



(11) **EP 0 477 972 B1**

K1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention  
of the grant of the patent:  
04.09.1996 **Bulletin 1996/36**

(51) Int Cl.<sup>6</sup> **C12Q 1/70, C12Q 1/68,**  
**C12P 19/34**  
**// C07H21/04**

(21) Application number: **91116561.1**

(22) Date of filing: **27.09.1991**

(54) **Nucleotide sequences useful as type-specific probes, PCR primers and LCR probes for the amplification and detection of human papilloma virus, and related kits and methods**

Nukleotid-Sequenzen nützlich als typenspezifische Sonden, PCR Primers und LCR Sonden zur Amplifikation und zum Nachweis von humanem Papillomavirus, sowie dazu verwendete Kits und Verfahren

Séquences nucléotidiques utiles comme sondes spécifiques du type amorces de PCR et sondes pour l'amplification et détection du virus-papilloma humain, et kits et procédés utilisés dans ce but

(84) Designated Contracting States:  
**AT BE CH DE DK ES FR GB GR IT LI NL SE**

(30) Priority: **28.09.1990 US 589948**  
**28.09.1990 US 590105**  
**28.09.1990 US 590253**

(43) Date of publication of application:  
**01.04.1992 Bulletin 1992/14**

(73) Proprietor: **ABBOTT LABORATORIES**  
**Abbott Park, Illinois 60064-3500 (US)**

(72) Inventors:  
• **Joseph, Jeffrey L.**  
**Cherry Hill, New Jersey 08002 (US)**  
• **Bouma, Stanley R.**  
**Mundelein, Illinois 60060 (US)**

• **Marshall, Ronald L.**  
**Zion, Illinois 60099 (US)**  
• **Laffler, Thomas G.**  
**Libertyville, Illinois 60048 (US)**

(74) Representative: **Modiano, Guido, Dr.-Ing. et al**  
**Modiano & Associati S.r.l.**  
**Via Meravigli, 16**  
**20123 Milano (IT)**

(56) References cited  
**EP-A- 0 301 968** **EP-A- 0 320 308**  
**EP-A- 0 402 132** **EP-A- 0 425 995**  
**WO-A-87/05630** **WO-A-89/05357**  
**WO-A-89/09940** **WO-A-90/02821**  
**WO-A-91/10675**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention)

## Description

This invention relates generally to human papilloma virus, and more particularly, relates to nucleotide sequences of short strands of human papilloma virus which can be amplified and/or used to determine the presence of human papilloma virus products in a test sample, and some of which also can be amplified and/or used to determine the specific type of human papilloma virus of types 16 and 18 present in the test sample

Human papilloma virus (HPV) is recognized as a venereally-transmitted disease of the anogenital tract which often is associated with the pathogenesis of cervical cancer and its precursor lesions. More than 55 types of HPV have been characterized. Of these, at least 21 types infect the anogenital tract. L. Gregoire et al., *J. Clin. Micro* 27 (12) 2660-2665 (1989). These mucosotropic viruses are associated most frequently with benign condyloma or latent infections. However, the presence of HPV in premalignant lesions and invasive cancers, particularly of the cervix, may reflect the oncogenic potential of these viruses. See P. M. Howley, in *Important Advances in Oncology*, D. T. DeVita, Jr. et al., eds., J. B. Lippincott, Philadelphia, PA (1987) at pages 55-73.

Certain HPV types, namely, HPV type 16 and type 18, and to a lesser extent HPV types 31, 33 and 35, are found in a high proportion of invasive cervical cancers and their metastases. However, many HPV types which infect the anogenital tract, such as HPV types 6 and 11, are found most commonly in benign condyloma and only rarely are found in invasive cancers. HPV detected in the anogenital tract can be classified broadly as low risk papilloma viruses (HPV types 6 and 11), intermediate risk papilloma viruses (HPV types 31, 33 and 35) or high risk papilloma viruses (HPV types 16 and 18), based on the association of the particular HPV type with malignancy. A. T. Lorincz et al., *J. Nat'l Cancer Inst.* 79 671 (1987). Thus, the detection of the presence of HPV and the determination of the specific type of HPV can provide a diagnostic and prognostic tool useful for determining the clinical significance associated with certain HPV types. The early detection of HPV by sensitive and specific reagents and methodologies also could provide earlier therapeutic management and counseling.

A need therefore exists for accurate and reliable methods to identify and type HPV in clinical specimens. However, known polyclonal antisera prepared by immunizing animals with disrupted virions are capable of detecting HPV antigens in only about 30-70% of cutaneous and mucosal warts. Further, the antisera are broadly cross-reactive. Available immunological tests have two major drawbacks. First, only well-differentiated cells apparently are capable of viral antigen expression. HPV-infected tissues which show higher degrees of neoplasia, such as carcinoma *in situ*, rarely contain HPV antigen. Thus, the further the development of the malignancy, the smaller the amount of detectable virus in the tested tissue. Secondly, these immunological tests are unable to identify specific viral types.

It is known that papilloma viruses share amino acid sequences in the major capsid proteins. See, for example, C. C. Baker, in *The Papovaviridae* (Vol. 2), P. M. Howley and N. P. Salzman, eds., Plenum Publ. Corp., New York (1987) at pages 321-385. The DNAs of this virus cross-hybridize, indicating homologous sequences. M. F. Law et al., *J. Virol.* 58 225-229 (1979). Thus, molecular hybridization techniques have been developed as a more sensitive and specific means of detecting and differentiating HPV DNA and RNA in clinical specimens. See A. T. Lorincz, *Obstetrics and Gynecol. Clinics of N. America* 14 451 (1987).

Sequences specific for the DNA and RNA of human papilloma virus are known and have been published. See, for example, PCT application No. WO 89/69940 published October 19, 1989, PCT application No. WO 86/05816 published October 9, 1986 and European Patent Application No. 0 301 968 published February 1, 1989.

The molecular hybridization techniques used to detect homologous DNA sequences are sensitive and can be highly specific if used with probes which bind to nucleic acid sequences which are unique to a particular HPV type. However, the concentration of total viral DNA in a given clinical sample may be below the limit of sensitivity of the test. For example, the amount of viral DNA in dysplastic cervical lesions is reduced with increasing dysplasia.

To overcome this problem of sensitivity, viral DNA sequences can be amplified by using, for example, the polymerase chain reaction (PCR) or the ligase chain reaction (LCR) techniques. The products thus obtained can be identified by using conventional hybridization techniques for identification of virus types, such as Southern blotting. See C. Ostle, *Biotechniques* 6:163 (1988). K. B. Mullis, U. S. Patent No. 4,683,202, and EP-A-320 308 (BioTechnica).

Both PCR and LCR serve to amplify the DNA present in a test sample to detectable levels. In practice, the level of sensitivity is about 50 to 100 copies per sample. The next most sensitive technique is dot-blot, which can detect about 10,000 molecules, while Southern blot reliably detects about 100,000 copies of DNA per sample.

Thus, the appropriate diagnosis of HPV may require two steps. In one strategy, the presence of a clinically relevant type of HPV is first detected with a group-specific primer. After the presence of HPV is detected, differentiation between types can be performed by using a type-specific probe having low homology between the HPVs of the group. Alternatively, differentiation can be performed using a mixture of type-specific probes at the outset, provided these probes amplify DNA independently of each other, and that they can be detected independently. In the past, such tasks were attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. The use of DNA-based tests increases both sensitivity and specificity over prior-art antibody-based tests.

It therefore would be advantageous to provide oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of HPV in a test sample. It also would be advantageous to provide short oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of specific types of HPV in the test sample. The combined use of oligonucleotide strands would be advantageous for allowing for the specific and sensitive *in vitro* diagnosis of the presence and specific type of HPV present in test samples.

#### SUMMARY OF THE INVENTION

Oligonucleotides of from about 10 to about 60 nucleotides are provided which can be amplified and used either to detect specific sequences of specific types of human papilloma virus, or consensus regions with high homology among different types. The presence of HPV is determined by contacting the test sample with sequences provided to detect the presence, if any, of HPV types 6, 11, 16, 18, 31, 33 and 61. This may be done with or without prior amplification, for example, by PCR or LCR. Either type-specific or consensus amplification is also possible. Two oligonucleotides are provided if the sequence is to be amplified by PCR, and four oligonucleotides provided if amplification is by LCR, in accordance with these known amplification procedures. After the presence of HPV is detected, the type of HPV present in the sample can be determined by using HPV type-specific probes, by subsequent rounds of PCR, or by LCR. Alternatively, the presence of type-specific HPV can be determined by contacting the test sample directly with type-specific nucleotide sequence provided by the invention for the detection of HPV types 16 and 18. Also provided are methods for using the oligonucleotides and kits for amplifying and detecting the presence of human papilloma virus.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1 and PCR5 were used to amplify selected plasmids wherein HPV 6 is in lane 1, HPV 11 is in lane 2, HPV 16 is in lane 3, HPV 18 is in lane 4, and HPV 31 is in lane 5, HPV 33 is in lane 6, HPV 61 is in lane 7, and molecular weight standards are in lane 8.

FIG. 2 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1, PCR2, PCR3, PCR4 and PCR5 were used to amplify plasmid p65 16.8 (HPV 16). PCR1 and PCR5 are primers according to the invention.

FIG. 3 is a photograph of the ethidium bromide-stained gels wherein PCR 1 4 and PCR15 are used in conjunction with IWDO to obtain amplified PCR product.

FIG. 4 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR5A, LCR5A', LCR5B and LCR5B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 5 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR6A, LCR6A', LCR6B and LCR6B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 6 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR7A, LCR7A', LCR7B and LCR7B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 7 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR8A, LCR8A', LCR8B and LCR8B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

#### DETAILED DESCRIPTION OF THE INVENTION

The appropriate diagnosis of HPV requires two sets of conditions. The first enables the detection of all pertinent types, and the second set allows differentiation among them. In the past, such tasks have been attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. Thus, the use of DNA-based tests tends to increase both sensitivity and specificity over antibody-based tests.

U. S. Patents No. 4,683,195 and 4,683,202 teach a method of amplifying DNA sequences by using PCR. This method now is a standard procedure in many molecular biology laboratories. Examples 1-3 which follow below utilize the procedures taught in these two patents and the method as described in the package insert of the commercially-available Gene-Amp™ kit (Document No. 55635-6/89, Perkin-Elmer/Cetus, Emeryville, CA).

In PCR, two complementary polynucleotide strands are amplified by treating the strands with two oligonucleotide primers such that an extension product of each primer is synthesized which is complementary to each nucleic acid strand. The primers are selected such that the extension product of one primer forms a template for the synthesis of an extension product from the other primer once the extension product of the one primer is separated from the template. A chain reaction is maintained by a cycle of denaturing the primer extension products from their templates, treating

the single-stranded molecule generated with the same primers to re-anneal, and allowing the primers to form further extension products. The cycle is repeated for any many times as it takes to increase the target nucleic acid segments to a concentration where they can be detected.

The amplified target sequence can be detected by any of several known techniques, for example, by denaturing the double-stranded products formed by PCR, and treating those products with one or more reporter probes which hybridize with the extension products. The reporter probe has a detectable label, and usually is added in excess. The unhybridized reporter probe, therefore, must be separated from the hybridized reporter probe by involving a separation step. In another method of detecting the extension products without reporter probe and a separation step, the extension products are detected by gels stained with ethidium bromide. The diagnosis can be confirmed by transferring the DNA to nitrocellulose and probing with a probe specific to the HPV type suspected of being present in the sample.

Alternately with PCR, one may take advantage of known restriction sites within the HPV DNA to demonstrate that the amplified DNA contains the expected sequence by examining the cleavage pattern(s) generated with one or more restriction endonucleases. Verifying the authenticity of the amplified sequence may be necessary for two reasons: (1) to ensure that sequences complementary to the amplifying primers are not fortuitously present in cellular DNA which does not contain HPV DNA, and (2), to identify the type of HPV present in the sample. If the sequences chosen for amplification are conserved among HPV types, then the finding of an amplified product does not implicate a particular HPV type. It also should be possible to predict the size of the amplified product based on the binding positions of the two primers. Thus, when that product is found, one reasonably can be assured that HPV is present. However, two different types of HPV may give the same or different size products. Thus, hybridization should be used to confirm the identity of the amplified sequence until confidence is built that the interpretation of the results is reliable. It should be pointed out that the PCR technique will identify only closely related, or type-specific sequences in the absence of highly homologous primers, since only a small portion of the genome is analyzed.

Another particularly useful detection technique is described in EP-A-357 011. In this method, a different reporter molecule, e.g. hapten, is attached to each primer. Following amplification, but before denaturation, duplexes can be detected by "capturing" one hapten (hapten1) with a solid phase coated with anti-hapten1. The separated complex can be detected with a conjugate of label and anti-hapten2, and label associated with the solid phase can be measured.

The Ligase Chain Reaction (LCR) amplifies sections of DNA by copying the section of DNA, and copying the copies of that section of DNA, many times over. This method is described in European Patent Application No. 0 320 308 published June 14, 1989, which is incorporated herein by reference. In this procedure, two probes (for example, A and B) complementary to immediately adjacent regions of a target sequence are hybridized and ligated. This ligated probe then is denatured away from the target, after which it is hybridized with two additional probes (A' and B') of sense opposite to the initial probes A and B. The secondary probes are themselves then ligated. Subsequent cycles of denaturation/hybridization/ligation create the formation of double-length probes of both sense (+) and antisense (-).

In LCR, the nucleic acid of the sample is provided either as single stranded DNA or as double-stranded DNA which is denatured to separate the strands. Four probes are utilized: the first two probes (A and B) are the so-called primary probes, and the second two probes (A' and B') are the so-called secondary probes. The first probe (A) is a single strand capable of hybridizing to a first segment of the primary strand of the target nucleotide sequence. The second probe (B) is capable of hybridizing to a second segment of the primary strand of the target nucleotide sequence. The 5' end of the first segment of the primary strand of the target is positioned relative to the 3' end of the second segment of the primary strand of the target to enable joining of the 3' end of the first probe to the 5' end of the second probe, when the probes are hybridized to the primary strand of the target nucleotide sequence. The third probe (A') is capable of hybridizing to the first probe, and the fourth probe (B') is capable of hybridizing to the second probe (B). The hybridized probes are ligated to form reorganized fused probe sequences. Then, the DNA in the sample is denatured to separate ligated probes from sample DNA. Successive cycles wherein the ligated probes and target DNA undergo the above-described process are performed to increase the amount of detectable DNA in the sample. The amount of cycles performed is dependent upon the sequence used and the sensitivity required of the test. Usually, the cycle can be repeated from 15 to 60 times. At least one of the probes can be conjugated to a signal generating compound.

If the four probes are conjugated to appropriate binding members, the detection of amplified product can be accomplished using standard manual or automated immunoassay procedures known to those skilled in the art. These procedures include, for example, immunochromatography, ELISA, EIA and MEIA. Hybridization also can be accomplished by following standard dot-, slot- or replica-blot procedures which are known to those in the art. The sequences can be labelled with an appropriate signal generating compound (label), which is capable of generating a measurable signal detectable by external means. The various signal generating compounds contemplated include chromogens, catalysts such as enzymes, luminescent compounds such as fluoroscein and rhodamine, chemiluminescent compounds, radioactive elements such as  $^{32}\text{P}$ , and other labels known to those of ordinary skill in the art. The selection of a particular label is not critical, but it will be capable of producing a signal either by itself or in conjunction with one or more additional substances. A variety of different indicator reagents can be formed of label and specific binding member. Either the label or a specific binding member can be varied. Examples of specific binding members which

can be used as a member of the indicator reagent include antibodies, both monoclonal, polyclonal, and fragments thereof; avidin or biotin, biotin and anti-biotin, a carbohydrate or a lectin, a complementary nucleotide sequence, an effector or a receptor molecule, an enzyme cofactor or an enzyme, an enzyme inhibitor or an enzyme, also any antigenic substances, haptens, antibodies, and combinations thereof

The test sample can be any biological material suspected of containing HPV. Thus, the test sample can be human body tissue, or a test sample which contains cells suspected of containing HPV.

The invention will now be described by way of Examples, which are meant to describe, but not to limit, the spirit and scope of the invention.

The following terms used in the examples are trademarks, tradenames or chemical abbreviations as specified

TRIS - chemical abbreviation for [tris(hydroxymethyl)aminomethane], used as a buffer.

EDTA - chemical abbreviation for ethylenediaminetetraacetic acid, a chelating agent.

FITC - chemical abbreviation for fluorescein isothiocyanate, a fluorescent hapten derivative.

NHS-ester - chemical abbreviation for N-hydroxysuccinamide ester

MES - chemical abbreviation for [2-(N-morpholino)ethanesulfonic acid], a buffer

TWEEN®-20 - trademark of Atlas Chemical for polyoxyethylene sorbitan monolaurate, a detergent.

BIS-TRIS - chemical abbreviation for [bis-(2-hydroxyethyl)-amino]tris-(hydroxymethyl)methane, a buffer.

TRITON X-100® - trademark of Rohm & Haas for nonaethylene glycol octylphenol ether, a detergent

IMx® - trademark of Abbott Laboratories for an automated instrument for performing microparticle enzyme immunoassay (MEIA).

## EXAMPLES

### EXAMPLE 1

PCR was performed essentially following the package insert of the commercially available Gene-Amp™ kit (document No. 55635-6/89, available from Perkin-Elmer/Cetus, Emeryville, CA). The following reagents were mixed in a 0.5 mL polypropylene tube and used in performing PCR.

Reagent	Final Concentration
Water	(to give final volume = 50 or 100 µL)
Reaction Buffer	10 mM TRIS pH 8.3 50 mM KCl 1.5 mM MgCl <sub>2</sub> 0.01% gelatin
dNTP mixture	200 µM each of dATP, dCTP, dGTP, and TTP
pCR1	1 µM
pCR2	1 µM
plasmid	10 µL 1 ng/100µL
(or control-human placental DNA (Pooled Placental DNA, catalog D-3287, Sigma Chemical Co, St. Louis MO).	
DNA polymerase, <u>Thermus Aquaticus</u>	25 or 63 9 units/1 mL

After mixing, the reaction mixture was overlaid with 100 µL of mineral oil. The tube then was placed in an instrument capable of incubation at several temperatures, and subjected to 30 or 40 cycles of programmed temperature change. The precise cycle of temperature change used, and the instrument used, varied with the experiment, and is detailed in the descriptions of the figures in Example 3.

### EXAMPLE 2

Following the procedure of Example 1, the following sequences were found to amplify sections of papilloma virus types 6, 11, 16, 18, 31, 33, and 61 using PCR.

PCR1: CAGATGTCTC TGTGGCGGCC TAGTG (ID No 1)

PCR5: AGGTGTCAGG AAAACCAAT TTATT (ID No 5)

PCR14: GAATTAGITA GACCATTTAA AAG (ID No 6)

PCR15: GGGGAAACAC CAGAATGGAT A (ID No 7)

IWDO ATCATATGCC CACTGTACCA T (ID No 8)

Sequence IWDO is derived from a sequence disclosed in international application number PCT/US86/00629 (WO 86/05816)

TABLE 1 shows the sequences and where they map to in the various types.

TABLE 1  
SEQUENCES WHICH CAN BE USED AS PROBES OR PCR PRIMERS

PROBE ID NO.	SEQ ID NO.	SEQUENCE	SENSE	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:
				(type 6)	(type 11)	(type 16)	(type 18)	(type 31)	(type 33)
PCR1:	1	CAGATGTCTCTGTGGCCGCTAGTG	+	5786-5810	5768-5792	5634-5658	5610-5634	5550-5574	5591-5615
PCR2:	2	CGTTTTCATATTTTTCGAGATG	+	5767-5791	5749-5773	615-5639	5591-5615	5531-5555	5572-5596
PCR3:	3	AAGTTGTAAGCACCAGTAATATGT	+	5844-5868	5826-5850	695-5719	5671-5695	5611-5635	5652-5676
PCR4:	4	AATGTACCTAATACCTATATTTG	-	6008-5984	5990-5966	865-5841	5841-5817	5784-5760	5825-5801
PCR5:	5	AGGTGTCAGGAAACCAATTTATT	-	6044-6020	6026-6002	5901-5877	5877-5853	5820-5796	5861-5837
PCR14:	6	GAATTAGTATGACCATTTAAAG	+	1495-1517	1495-1517	1524-1546	1595-1617	1462-1484	1518-1540
PCR15:	7	GGGAAACACCAAGATGGATA	+	1834-1854	1834-1854	1863-1883	1934-1954	1801-1821	1857-1877
SIWDO:	8	ATCATATGCCACTGTACCAT	-	1931-1911	1931-1911	1960-1940	2031-2011	1898-1878	1954-1934

note: PCR2, PCR3 and PCR4 are not probes or PCR primers of the invention

#### EXAMPLE 3

Linearized plasmids containing full-length papilloma virus inserts in pGEM3 were used as targets. These were pHPV6.1 (HPV6), pSP65.11.5 (HPV 11), p65.16.8 (HPV16), pHPV18H(HPV18), pG3 HPV31 (HPV31), pLNK322,HPV33 (HPV33), and pBR322 HPV61 (HPV61). The Programmable Cyclic Reactor™ (available from Ericomp, San Diego) was used as the incubation instrument. Following PCR procedures as described in Example 1, 10 µL aliquots were analyzed by electrophoresis through agarose (comprising a 5:3 ratio of NuSieve® SeaKem® GTG, available from the FMC Corp., Rockland, ME) in a buffer comprising 0.089 M TRIS, 0.089 M borate, 2 mM EDTA, and 0.5 ppt ethidium bromide.

FIG 1 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using 63.9 units/mL DNA polymerase, in the DNA Thermal Cycler™ (Perkin-Elmer/CETUS, Emeryville, CA). The samples were heated for 5 minutes at 94°C, then subjected to 40 cycles of a temperature program of: 1 minute at 94°C, 2 minutes at 40°C, and 1.5 minutes at 72°C. The PCR primers used in this case were PCR1 and PCR5 of Example 2. Examination of the gel following electrophoresis showed bands at the expected positions, i.e. 292 bp. Lane 1: HPV6; lane 2: HPV 11; lane 3: HPV16; lane 4: HPV 18; lane 5: HPV31; lane 6: HPV33; lane 7: HPV61; lane 8: pooled human placental DNA (suspected of having HPV infection); lane 9: molecular weight markers-Hae III digest of ΦX174.

FIG 2 is a photograph of an ethidium bromide-stained 4% agarose gel showing results using 25 units/mL DNA polymerase, in the Programmable Cyclic Reactor™ (Ericomp, San Diego, CA). Samples in this case were subjected to 30 cycles of a temperature program of: 50°C for one (1) minute, 72°C for two (2) minutes and 95°C for one (1) minute. In this case, the primers PCR1, PCR2, PCR3, PCR4 and PCR5 of Example 2 were used to amplify plasmid

p65,16,8(HPV 16). Examination of the gel of Figure 2 shows bands at the expected positions, i.e. PCR 1 and PCR4, 235 bp, lane 2; PCR1 and PCR5, 267 bp, lane 4; PCR2 and PCR4, 254 bp, lane 6; PCR2 and PCR5, 266 bp, lane 8; PCR3 and PCR4, 174 bp, lane 10; PCR3 and PCR5, 206 bp, lane 12; molecular weight marker, 123, 246, 369, 492, ... bp ladder, lane 1. Note footnote to Table 1

FIG. 3 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using the same conditions as FIG. 1. In this case, PCR14 and PCR15 were used as primers in conjunction with IWDO. The expected size of the amplified PCR product of PCR 14 and IWDO is 437 bp for all of the HPV types tested. The expected size of the product of PCR 15 and IWDO is 98 bp. Products of these sizes appear in the gels, confirming that PCR14 and PCR15, used in conjunction with IWDO, will amplify HPV DNA of types 6, 11, 16, 18, 31, 33, and 61. Lane 1, Molecular weight marker (Hae III digest of FX 174). PCR 14 + IWDO, lanes 2-9, lane 2, HPV6, lane 3, HPV 11, lane 4, HPV16, lane 5, HPV18, lane 6, HPV31, lane 7, HPV33, lane 8, HPV61, lane 9, human placental DNA suspected of being infected with HPV; PCR 5 + IWDO, lanes 10-17, lane 10, HPV6, lane 11, HPV 11, lane 12, HPV16, lane 13, HPV18, lane 14, HPV31, lane 15, HPV33, lane 16, HPV61, lane 17, human placental DNA suspected of being infected with HPV, lane 18, molecular weight marker (Hae III digest of FX174 and Hind III digest of 1 DNA).

#### EXAMPLE 4

The following reagents were mixed in a 0.5 mL polypropylene tube as follows for the Ligase Chain Reaction (LCR)

Reagent	Volume	Final Concentration
Water	21 $\mu$ L	
Reaction Buffer	10 $\mu$ L	50 mM EPPS pH7.8 10 mM $\text{NH}_4\text{Cl}$ 10 mM $\text{MgCl}_2$  100 mM $\text{K}^+$ (from all sources) 0.001% BSA 1 mM DDT
Nicotine Adenine Dinucleotide (NAD)	0.5 $\mu$ L	100 $\mu$ L
ProbeA (sense)	4 $\mu$ L	$5.0 \times 10^{11}$ molecules
ProbeA' (antisense, 5'-phosphate)	4 $\mu$ L	$7.5 \times 10^{11}$ molecules
ProbeB (sense, 5'-phosphate)	4 $\mu$ L	$7.5 \times 10^{11}$ molecules
Probe B' (antisense)	4 $\mu$ L	$5.0 \times 10^{11}$ molecules
Target (including human placental carrier DNA at 10 $\mu$ g/mL)	1.5 $\mu$ L	15 ng/50 $\mu$ L
DNA ligase, <i>Thermus thermophilus</i>	1 $\mu$ L	

This reaction mixture was overlaid with 30  $\mu$ L of mineral oil. The tube was placed in an instrument capable of incubation at several temperatures (e.g. thermal cycler from Coy Laboratory Products (Ann Arbor, MI) or the Programmable Cycler Reactor™ (available from Ericomp, San Diego, CA), and then subjected to several cycles of programmed temperature change. Each cycle involved incubation at 50°C for one minute and 85°C for one minute.

#### EXAMPLE 5

The following procedure was used when performing the Ligase Chain Reaction (LCR), which is described in published European Patent Application No. 0 320 308 A2. The reagents of Example 4 were utilized in the procedure as follows. Two probes (A and B) complementary to immediately adjacent to regions of a target sequence were hybridized and ligated. This ligated probe was denatured away from the target, and hybridized with two additional probes (A' and B') of sense opposite to the initial probes (A and B). The secondary probes then were ligated. Subsequent cycles of denaturation/hybridization/ligation created the formation of double-length probes of both + and - sense.

#### EXAMPLE 6

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16

<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to</u>
LCRSA	81	GCTGCAACA ACTATACATG ATATAA	157 - 182
LCRSA'	82	TTATATCATG TATAGTTGTT TGCAGC	182 - 157
LCRSB	83	TTATTAGAATG TGTGTACTGC AAGCA	183 - 208
LCRSB'	84	TGCTTGCACT ACACACATTC TAATA	208 - 157

## EXAMPLE 9

Base-denatured plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. These plasmids were pG3HPV6(+) (HPV6), pSP 65, 11 5 (HPV11), pSP65 168 (HPV16), p63HPV18H(-) (HPV18), p63HPV31 (HPV31), pLNK322 HPV33 (HPV33), pBR322 HPV35 (HPV35), pUC19 HPV52 (HPV52), pLNK322 HPV58 (HPV58), pUC9 HPV59 (HPV59) and pBR322 HPV61 (HPV61). All of the oligonucleotides used as probes from Example 8 had chemical labels covalently attached at the ends distal from ligation. These labels were: 5'-fluorescein-LCRSA, 3'-fluorescein-LCRSA', 3'-biotin-LCRSB and 5'-biotin-LCRSB'. Covalent attachment was performed by known methods, i.e., reaction of amine-terminated oligonucleotides with FITC or biotin-NHS-ester essentially following the procedures of Kansal et al. Tel. Letters 29:5537-5540 (1988). The thermal cycler used was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed using a prototype version of the IM<sub>2</sub> instrument (Abbott Laboratories, Abbott Park, IL), following the protocol for microparticle enzyme immunoassays as follows: A 40 µL aliquot of an LCR mixture was diluted 1:1 with distilled water. This diluted mixture was incubated with 50 µL anti-fluorescein-conjugated polystyrene microparticles for five (5) minutes to form a suspension of immune complexes on the microparticles. This suspension then was transferred to an inert glass fiber matrix, to which the microparticles became attached. The matrix was washed with buffer (0.3M NaCl, 10 mM TRIS pH8, 0.1%Na<sub>2</sub>S<sub>2</sub>O<sub>5</sub>). Any immune complexes attached to the glass matrix was detected by using alkaline phosphatase-labeled conjugate that catalyzed the hydrolysis of 4-methylumbelliferone. The rate at which the 4-methylumbelliferone was generated on the matrix was proportional to the concentration of LCR product formed in the reaction mixture.

Referring to FIG. 4, the graph shows the results obtained from performing LCR on 10<sup>7</sup> molecules of the targets in shown. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV16, and those values are about 60 times background signal.

## EXAMPLE 10

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18.

<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to</u>
LCR6A	85	CTTCACTGCA AGACATACAA ATAA	172 - 195
LCR6A'	86	TTATTTCTAT GTCTTGCACT GAA	195 - 173
LCR6B	87	TCCTGTGTATA TTGCAAGACA GTAT	196 - 219
LCR6B'	88	TACTGTCTTG CAATATACAC AGG	218 - 196

## EXAMPLE 11

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids used were those described in Example 9. All of the oligonucleotides used as probes obtained from Example 10 had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following LCR procedure described in Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IM<sub>2</sub> instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 5, the graph displays the results obtained from performing LCR on 10<sup>7</sup> molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second.



second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 8, and those values are about 40 times background signal

#### 5 EXAMPLE 12

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18

10	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR7A	89	TATATTGCAA GACAGTATTG GAAC	200 - 223
	LCR7A'	90	pGTTCCAATAC TGTCTTGGAA TTTA	223 - 200
	LCR7B	91	pTTACAGAAGT ATTGAATTT GCATT	224 - 249
15	LCR7B'	92	AATGCAATT CAAATACCTC TGTA	249 - 224

#### EXAMPLE 13

20 Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids were those of Example 9. All of the oligonucleotides from Example 12 which were used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was as described in Example 11.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

25 Referring to FIG. 6, the graph shows the results obtained from performing LCR on  $10^7$  molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 18, and those values are about 80 times background signal.

30

#### EXAMPLE 14

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16

35	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR8A	93	GTATGGAACA ACATTAGAAC AGCA	352 - 375
	LCR8A'	94	pTGCTGTTCTA ATGTTGTTCC ATAC	375 - 352
40	LCR8B	95	pATACAACAAA CC6TTGTGTG ATTT	376 - 399
	LCR8B'	96	AAATCACACA ACGGTTTGTG GTAT	399 - 376

#### 45 EXAMPLE 15

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. All of the oligonucleotides from Example 14 used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was as described in Example 11.

50 Following LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 7, the graph details the results obtained from performing LCR on  $10^7$  molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 16, and those values are about 36 times background signal.

55

EXAMPLE 16

The attached Appendix (example 16) discloses the sequences of the invention aligned to known sequences

EXAMPLE 16

APPENDIX

HUMAN PAPILLOMA VIRUS

ALIGNMENT of TYPES 6, 11, 16, 18, 31, and 33, with CONSENSUS SEQUENCE

The appendix lists the sequences of HPV types 6, 11, 16, 18, 31, and 33. It also shows where the sequences of this invention line up with respect to these HPV sequences. In addition, the appendix shows where other sequences, known to the inventors as of 28 September 1990, and claimed or disclosed by or unknown to others, line up with respect to these sequences.

1. Sequences and Regions Claimed by Us;

PCR = Sequences per examples 1 through 3 (only PCR1, PCR5, PCR14 and PCR15)

LCR = Sequences per examples 4 through 14 only

2. Sequences and Regions Unknown to Others and Not Claimed by Us;

PCR = Sequences designated PCR other than those above JJ

LCR = Sequences designated LCR other than those above

3. Sequences and Regions Claimed by Others;  
(*Italics represents antisense sequences*)

AUS = International application number (Australians) PCT/AU88/00047 (WO 88/06634)

WL = International application number (Wayne Lancaster, Wayne State University) PCT/US86/00629  
(WO 86/05816)

BE = European Patent Application (Belgians) 89.033834 (X = T or U)

C = International application number (CETUS) PCT/US89/03747 (WO 90/C2821)

O = International application number (Oncor) PCT/US89/O1318 (WO 89/09940)

and

4. Sequences and Regions Disclosed by Others.

S = Sarkar, F.H. and Crissman, J.D. *Biotechniques* 9:180-184 (1990) (*Italics represents antisense sequences*)

6 1 gTTAATAACAATCTTgGTTTAA AAAAAGGAGGG ACCGAAA ACCGGTTCAACCGAAAA  
 11 1 cTTAATAACAATCTTAGTTTAA AAAAGAGGAGGG ACCGAAA ACCGGTTCAACCGAAAA  
 5 33 1 gtaAaCTATATgCCaAGTTTAA AAA AGtAGGCGTAACCGAAA gCGGTTCaACCGAAAA  
 16 1 actACAATAAT tcatGTATA AAA ctaAGGGGCTAACCGAAA cCGGTTCaACCGAAAc  
 31 1 TAATA ATAATAAT ctTAGTATA AAA AAgTAGGGAGTgACCGAAA GtGtGTAACCGAAAA  
 10 18 1 attAATActTTtaAcaattgTAGTATATAAA AA AGGGAGTAACCGAAAAGcGtGgGACCGAAAA  
 con --taataata-ta-aa-tcttag-T-tA-AAAaaag-ACGGAgtaACCGAAA-acggtt-aACCGAAAA  
 C4-GCCAAATGCGTTTT  
 C5-GCCAGCCTGGCTTTT  
 C16-CGGTTSAACCGAAAA  
 C17-CGGTCGGGACCGAAAA  
 C18-CGGTTSAACCGAAAH  
 C19-CGGTTCAACCGAAAH  
 015-ATTAATACTTTTAACAATTGTAGTATATAAA AA AGGGAGTAACCGAAAACGGTCGGGACCGAAAA-015  
 024- ACTACAATAAT TCATGTATA AAA CTAAGGCGGTAAACCGAAA TCGGTTCaACCGAAAC-024  
 20 S1-CGGTCGGGACCGAAAA  
 S3-ACCGAAAC  
 6 58 OGGTTgTATATAAA CCAGCCCTAAAaATTAGCAaACGAGCATTATGGAAGTgcAAATGCCCTCCAC  
 11 58 OGGTTaTATATAAA CCAGCCCAAAAAATTAGCAGACGAGGCATTATGGAAGTAAAGATGCCCTCCAC  
 25 33 62 CGGTgcATATATAAGCA aNCATTTTgcagtaAGtActGCACgActATGTTTCAAGCacTgAGGA  
 16 58 OGGTTaGTATA AAAGCA gACATTTTatGcaCcaAAAgAGAACTcGcaATGTTCCAcAGGA  
 31 60 OGGTTgGTATATAAAGCAcetaGTATTTgtGcaAAccTACAgacGCcATGTTcaAAaATCCTgCAGA  
 30 18 66 CGGT GTATATAAA agatgtGagaacacacCcAcaATACtatgGCgcgtTtAggATCcaACgCg  
 con OGGTt-gtatataAagcag--ca-a--at--gcaaaac-aggatt-cgatgttt-aagAtcC--c-ga  
 GCC-C4 AUS1-ATGCCCTCCAC  
 GCC-C5  
 CGG-C16 C67-AAATCCTGCAGCA  
 CGG-C17 C68-CCTACAGACGCCATGTTCa-C68  
 CGG-C18 C71-GCAGTAAGGTACTGCAC-C71  
 CGG-C19 010-OGATCCCAACAG  
 015-CGGT GTATATAAA AGATGTGAGAAACACACCAATACTATGGCGGCTTTGAGGATCC-015  
 024-CGGTAGTATA AAAGCA GACATTTTATGCCAACAAAGAGAGACTGCAATGTTTCCAGAGAI-024  
 40 CGGTG-S1 S2-CCGCGGGAATCTCTAGGTTGTGC-S2  
 CGGTGTAGTATA AAAGC-S3

6 125 GTCTGCAACgCcAtAGACCAgTGTGCAAGAGCTTTAACTcTcTATGCAcAGgtTCAAAATTaAT  
 11 125 GTCTGCAACAcTcTATAGACCAgTGTGCAAGAGCTTTAACTcTcTtTGCAcAcCtTGCAAAATTCAGT  
 5 33 129 aaaACCAcGAACATgGcAtgAtTTGTGCAAGCAtTGAgACAACATATACACAACAtTgAAcTACAGT  
 16 124 gcGACCCcAGAAgTtAcacAgTTATGCAcAGAgcTGCcAAACAACATATACATGAAtATATATTAGAAAT  
 31 128 aaGACCTcGgAAATtGcATGAACtAAcGCTCGGcAtTGgAAATAcCctacgATGAAcTAgATTgAAAT  
 10 18 131 gcGACCcTcAcAAgctAcCtGATcTgtGCAcGGAcTGAACAcCtCactgcAcAGAcTAgAAATaAcCT  
 con g-gacCaagaa--tTacat-AgtTgtGCA-ggc-tTgAA-a-atCtatgcAt-a-aTa-aAaTaaa-T  
 GTCTGCAAC-AUS1 AUS7-GCAAGACGTTAAATCT-AUS7 C74-ACACTCTGCAAAATTCAGT  
 AAGACCTC-C67  
 15 010-GCGACCCCTACAAGCTACCTGATCTGTGCAAGGAACtGAACACTTCACtGCAAGACATAGAAATTAACCT-010  
 024-GCGACCcAGAAAGTTACCAcAGTATGCAcAGAGCTGCAAAcAACTATACATGATATATAATTAGAAT-024  
 S4-CTGGGCTCTTCAATGGTGTCAATA-S4  
 6 193 CctGTGTTTTCGAaGAATGCACtGACCAcAGcAGAGATTtATcCATATGcATATAAcACCTAAAGGTC  
 11 193 GCGTGTtTTTCAGgAATGCACtGACCAcAGcAGAGATATATGcATATGcTATATAgAcACCTAAAGGTT  
 20 33 197 GCGTgAAAGCAaAAACcTTGCAaAGcAGcTGAAGTATATGATTTGcATTtGcAGATTtAAcAGTT  
 16 192 GTGTGTACTGCAAGcAACAGTTACTgCAcgtGAGGTATATGAcTTTGcTTTCgGATTtATgCAATA  
 31 196 GTGTcTACTGCAaaggTcAGTTAAcAgcAAcAGAGGTATTGAATTTCGATTtACAGATTtAAcCAATA  
 25 18 199 GTGTaTACTGCAAgacagtaTtggaaCtTACAGAGGTATTtGaaTTTGcATTtAAAGATTtATTtTgTg  
 con GTGTgtatTGAaagaa--catTgacac-a-caGAGTtATgaaTtTGCAcTtAaaqAttTAA--gt-  
 AUS2-TACGTGACTGGTGCCCTGTC-AUS2 C73-ACACCTAAAGGTC  
 GC-C74 AUS3-TGAGGTATATGCACTTTTGCTTTT-AUS3  
 30 C69-GAGGTATTTGANTTTGC-C69 01-CTAAAGGTT  
 C61-GAGATWTATKCATATGC-C61 02-CTAAAGGTT  
 C69-ACAGTATTGGAACCTTACAG-C69 04-GATTTCCAA  
 C70-CAACAGTTACTGCGAGC-C70 06-TTATGCATA  
 C72-GACAGTATTGGAACCTTACAG-C70 07-TTATGCATA  
 S5-GTGTTTTTGCAGGAATGCACTGACCA-S5 08-AAATCGTAT  
 35 010-GTGZATATTGCAAGACAGTATTGCAACTTACAGAGGTATTtGAATTtGCATTtAAAGATTtATTtTGTG-010  
 011-TTATTGTG  
 012-TTATTGTG  
 013-AATAAACAC  
 017-CTAAAGGTC  
 018-CTAAAGGTC  
 40 024-GTGCTTACTGCAAGcAACAGTtACTGCGAGCTGAGGTATATGACTTTGCTTTTCGGGATTtATGCATA-024  
 025-TTATTGTG

```

6 261 cTGTtctGAGGcGgGtATtCCatATtCCAGCcTgGcGtTgTGCcTAGAArttCAtGGaAAAAATaAACCA
11 261 GTGTgGcGAGAcAactTtTCCcTTTGCAGCgTGTGccTGTtGCCTAGAAcTgCAAGGgAAAAATTAACCA
33 265 GTATATAGAGAGcGgAAATCCATTTGgAAATGTaAactgTGTtTgcgGTtTcTATCTAAAAATTAGTGA
16 260 GTATATAGAGAGAGcGgAAATCCATATGctGTATGTgATAAAATGTTTAAAGTTTTATCTAAAAATTAGTGA
10 264 GTATATAGgGACGacACACACACcGgAGTgTGTGacAAAAATGTTTAAgATTTTATTCAAAAgTAAgTGA
18 267 GTGTATAGAGAcagTATACCCcATGctGcATGccATAAATGTATAgATTTTATTCtagAAATTAGAGa

con GT-TatAGAGacggcAatCCatAtGcag-aTGTg--aaATGttTagaattTcattctAaAaATTAgTgA
15 C-44 CTCCTGTCGWHAGGTAWACSH-C44 JJ1-sattagnga
C-45 CTCCTGTCATATGCGGTACGA-C45 AUSB-GTGA
C-46 CCTGCTGCTGTGCTGTGCTT-C46 S6-GT
C-47 CYCTGCTGWHAGGTAWACSH-C47
C-48 CYCTGYGWHAGGTAWACSH-C48
C-49 CYCTGYGWHAGGTAWACSH-C49
20 C56-HGAGACRGCHWTCCATWTG-C56
C57-HGAGACRGSHWTCCATWTG-C57
C58-HGAGACRGVWWTCCATWTG-C58
C59-AGAGACAGTATACCCCATG-C59
GTGTGCGGAGACACTTTCCCTTTGCAGCGTGTGCGCTGTG-01
GTGTGCGGAGACACTTTCCCT-02
25 03-CAACTTTCCCTTTGCAGCGTGTGCGCTGTG-03
CACACCGCTCTGTGTAAGGGGAAACGTGCGACACGGACAC-04
GTATATAGAGATGGAATCA-06
GTATATAGAGATGGAATCAATCTGTATGTATGAATG-07
CATATATCTCTACCCCTTAGGTATACGACATACACTATTTAC-08
09-ACCCCTTAGGTATACGACATACACTATTTAC-09
30 010-GTGTATAGAGACAGTATACCCCATGCTGCATGCCATAAATGTATAGATTTTATTCAGAAATTAGAGA-010
GTGTATAGAGACAGTATACCC-011
GTGTATAGAGACAGTATACCCCATGCTGCATGCCATAAATG-012
CACATATCTCTGTATATGCGGTACGACATCAAGTATTTAC-013
014-GTATATAGAGATGGAATCAATGCTGTATGTATGAATGTTTAAAGTTTTATTCAAAAATTAGTGA-024
017-CTGTTTCGAGGCGGCTATCCA-017
35 018-CTGTTTCGAGGCGGCTATCCATATGCAAGCTGCGCGTGTG-018
019-GCCGATAGGTATACGTCGGACGGCACGAC-019
GACAAAGCTCCGCGGATAGGTATACGTCGGACGGCACGAC-020
024-GTATATAGAGATGGAATCAATGCTGTATGTATGAATGTTTAAAGTTTTATTCAAAAATTAGTGA-024
GTGTATAGAGACAGTATACCC-025
026-CAGTATACCCCATGCTGCATGCCATAAATG-026

```



```

6 464 aaccAAGGCGCGTTCATAAA          gCTAAAtgtacGTGGAAGGG          TCGGTG
      |||||
11 464 gggAAAGGCaCGTTCATAAA          CTAATaaCcaGTGGAAGGG          TCGTTG
      |||||
13 469 ttAAAcAAaGATTCATATAT          TtcGGGTGgtTGGGCAAGGCGGTGTgtgcGcgTGTG
      |||||
16 464 AAAAAGcAAaGATTCCATAATATA          aGGGTGCGTGGACGGTcGATGTATgtctTGTG
      |||||
10 31 468 AAAAAGaAACGATTCCACAACATAG          GaGGAAGTGGACaGGAcGtTGCATgCatGTG
      |||||
18 471 gAAAAcgaACGATTcACAACATAGctgggcactataGagGccaGtgccattcgTGctgcaaccGagc
con  aaaaAa--acgatTtCatAa-atag-----ctasaaggacg-tgGgcagggcg-tgcattggt-Gttg
      TGGTGATAGA-AUS9          AUS6-AAATGTATAGATTTTATTTC-AUS6
15 010-GAAAAACGACGATTTCCACAACATAGctGGGCACTATAGAGGCCAGTGCCTTCGTCTGCAACCGAGC-C65-CAACCGAGC
024-AAAAAGCAAGATTTCCATAATATA          AGGGGTCTGGTGGACCGGTGATGTATGTCTGTGTG-024

6 512          CcTACACTGC          TGGACAACATGCATG          GAAGACATGT
      |||||
20 11 512          CttACACTGC          TGGACAACATGCATG          GAAGACtTGT
      |||||
33 528          gaggtcccgACGTAGAGAACTGCactgtgAcgTGTAaaaaGcgCATGagagGACACaagcC
      |||||
16 523          cagatcaccAAGNaCACCTAGAGAAAC          CCagctGTAA          tCATGCATGGAGatACAC          C
      |||||
25 31 527          GagAAGACCTcGTactGAAAC          CCAagTGTAA          cATGCGTGGAGaAACAC          C
      |||||
18 539          acgacaGgaACGACTcCaacgacgcAgagaacacaCaAgtataAtattAaCtaTgcAtggACtcaaggC
con  --ga--gagaagaccacgta-aga-Actgca---ccaggtgtAaaacatgcaTGCgagagAcacaaagc
      C64-GAACACTAGAGAAAC          CCAG-C64
30 010-ACGACAGGAACGACTCCaACGACGCAAGAGAAACACAAAGTATAATATTAAATGATGCATGGACCTAAGGC-010
024-CAGATCATCAAGAACACGTAGAGAAAC          CCAG-024

6 547 TACCCTAAAGGA          TATTGTAtTAGACCTGCAaCCTCCaGACcCTGTAGGGTTACATTGCTATG
      |||||
35 11 547 TACCCTAAAGGA          TATcGTactTAGACCTGCAgCCTCTGACCCGTAGGGTTACATTGCTATG
      |||||
13 590 aACgTTAAAGGA          ATATGtTtTAGA          TTTatATCCTGAaCCAACTGacCTATACtGCTATG
      |||||
16 579 TAcAtTGCatGA          ATAtaTGTtTAGA          TTTGCAACCaGAGaCAaCTGAtcTCTACtGTtATG
      |||||
40 31 577 TAcgtTGCaAGAC          TAtgtTtTAGA          TTTGCAACCTcGAGcCAaCTGACCTCCACGTtTATG
      |||||
18 607 aAcAtTGCaAGAcattgtatTgcattTAGAgccccaaaAtgsaattcCggtTGACCTtCcatGTcAcG
con  tAc-tT--AgGAc-----at-tgt-tTAGAcctt---cacc-c-ga-cCa--tGaccTaccatG-tAGt
      BE16-ACCAGAGACAACXGAXCXAXCAXG-BE16
      BE18-GXAXGAXXXGCAACCGAGACAACXGAXCXAXC-BE18
010-AACATTGCAAGACATTGTATTGCATTAGAGCCCCAAATGAATTCGGTGTGACCTCTCATGTGACG-010
      C89-G
      C90-G
50
55

```





6 728 CTGGTTGTGcAGTGCaCAGaaacAGACATCAGAgAAgTgCAaCagCTTcTGTGGGcACATCAAAcAT  
 11 728 CTGGTTGTGgAGTGCcACAGAcgGAGACATCAGAgAAcTACAagAcCTTcTGTGGGcACATCAAAATAT  
 33 771 TtATGTTGTCaAcACTcACAGcAaGtGACcTcAGAACcATACAgAaCTacTtATGGGcACAgTgAAATAT  
 16 760 TTGTGcGTACaAAGcACACAcCTAGcATTCGcAcTTGgAAGAcCTGTAAATGGGcACAcTAgGAAAT  
 31 758 TTGTGTTGTAAGcACCAcCAAGTAgATATTTCGcATATTGCAAGAGCTGTAAATGGGcCTcATtTGGAAAT  
 18 809 cTagtaGTAGAAgCTcAGcAGAcGAcCTTCGAgcATTcCAGcAGCTGTTtTgAAcAcCccTgtccT  
 con -Tg--tGTAcAGaGcaCgAag--aGAcA7TcGacatTgcAa-AgCTgtT--aTgggcaCacTaaa--aT  
 XGc-BE19 BE29-AGCAAGXGACCCXACGAACCAKACA-BE29 C42-CCCCTGTGAYYYDTA  
 XKXKCCXKAC--BE20 C43-CTTGTGGGACAGGA  
 CXAGX-BE25  
 15 BE30-AGKACAGCAAGXGACCCXAGAACCAKACAGCAACX-BE30  
 010-CTAGTACTAGAAgCTcCAGcAGAcGACCTTCGAGCATCCAGCAGCTGTTTCTGNAcACCCCTGTCCTT-010  
 6 796 aGTGTGTGCCATCTGCGC AC CgAAgaCcTAACAaGATGGCGGAGGATTCAGGTACAGAAAAT  
 11 796 TGTGTGTGCCATCTGCGC AC CaAAACaCTAACaAGATGGCGGAGGATTCAGGTACAGAAAAT  
 33 839 TGTGTGCCcTAcCTGTGC ACaAcAAaTAACATCATcTAcAaATGGCcGATcCTGAAGGTACAAATGgg  
 16 828 TGTGTGCCcATcTGTTCCT CAgAAAcCaTAATCTAcCAtGGCTGATCCTGCAGGTACcAATGGGGAA  
 31 826 cGTGTGCCcCAAcTGTTCCT aCTAgAcGTAA CTACAAATGGCTGATTCAGCAGGTACAGAGGGGA  
 25 877 TGTGTGCCgtgTGTGc atCceagCagTAAGcACAAATGGCTGATTCAGAGGTACAGAGGGGA  
 con tGTGGT-CCcAtcTGTgTcaca--aaacaataatcaCAATg---G-t---g---g---ta-ag-ggat  
 C40-CACACRGGGTAGACRGG-C40 C75-ATGGCKGAYCCTGHAGGTAC-C75  
 C41-CACACAGGcACCACAGC-C41 C76-ATGGCKGAYGATTcAGGTAC-C76  
 ACACAC-C42 C77-ATGGCKGAYCCTTCAGGTAC-C77  
 30 ACACAC-C43 C81-TACCGMCTRGACCTCCATG-C81  
 C82-TACCGMCTRGCTAAGTCCATG-C82  
 C83-TACCGMCTRGGAAGTCCATG-C83  
 010-TGTGTGTCCGTGTGTGC ATCCcAGCAGTAAGCAACAATGGCTGATC-010  
 36 6 859 GAGGGGTcTGGGTGTACAGGATGGTTTATGGTAGAAGCTATAGTgcAaCACCcAaCAGG TAC  
 11 859 GAGGGGTcTGGGTGTACAGGATGGTTTATGGTAGAAGCcATAGTAGAGCACCACAGG TAC  
 33 906 GcTGGGATGGGTGTACTGGTGTGTTTgAGGTAGAAGcATcATAGAGAgAgaAACAGG aGA  
 16 895 GAGGGcACGGGATGTAATGGA TGTTTATGTAGAGcCTcGTAgTgGaaAaAAAAACAGG GGA  
 40 31 891 GGGcACGGGATGCAATGG TGTGTTTATGTAGAAGCAGTAaTcGACAgAcAgACAGG GGA  
 18 943 GGGcACGGGTGTAaAGGGGTGTTTATGTAGAAGCAGTgTgTAgCAaAaAaACAGGagatgtaat  
 con gagGGgAcgGGGTGtA-tGGA TGTTTtA-GTAGAAGCT-TagTagA-aaaaaACAGG-----a  
 C78-TGTANHGG:ITGGTTTATGT-C78  
 C79-TGTANHGG:ITGGTTTACAGGT-C79  
 C80-TGTANHGG:ITGGTTTATGGT-C80  
 C84-ACATKXCCXACCAAAATACA-C84  
 C85-ACATKXCCXACCAAAATACA-C85  
 50 C86-ACATKXCCXACCAAAATACA-C86  
 55

```

6 921 ACAAAATATCAGACGATGAGGAtGAGGAGGTGAGGACAGTGGGTATGACATGTGGACTTTATTGATG
11 921 ACAAAATATCAGAGATGAGGAAaGAGGAGGTGGAGGACAGTGGGTATGACATGTGGACTTTATTGATG
33 968 TAAATATCTCAGAGATGAGGATGAAaCaGcAGTACAGTGGcagGATTtAcTAGAgTTTATAGATG
16 957 TgcTATaTCAGATGACGAGAAcGAAAAATGacAGTGATACaGGtGaaGATTtGGTAGATTTTATAGtAA
10 951 caacATTTCAGAGGACGAAATGAAaGACAGcAGTGATACtGGGAGGATATGGTtGAcTTTATAGcAA
18 1009 atcagaTgacGAGGACGAAAAATG caACAG AcACaGGGtcGGATATGCTaGATTTTATTGATa
con a-aast-tcaGA-GA-GAg-AtGaa-a-g-ggatgAcA-tGGgtagGataTggTaGAcTTTATtGat-

15 6 989 A CAGcaATATTACA CACaATTcAcTGGaAGCACAGGCATTGTTTTAcAGGCAGGAGGCG
11 989 A CAGgcATATTACA CAaAAATCTGTGGaAGCACAGGCATTGTTTTAcAGGCAGGAGGCG
33 1036 ATtCTATgGAAaATAgTATACAGGCAGAcACAGAGGCAGCcCgGGcATTGTTTTAAATaACAGGAAGG
20 16 1025 ATgaTAATGATtATttaaACAGGCAGAAACAGAGACAGCACATGcgtTGTtTAcTCCACAGGAAGCa
31 1019 ATtGTAATGtATAcacaaAAtCAGGCAGAAgCAGAGACAGCACAGGCATTGTTTATGcACAGGAAGCg
18 1071 cacaaggaacATtttgtgAAcAGGCAGAgcTAGAGACAGCACAGGCATTGTTTATGcCAGGAgGtc
25 con attataatgcataataataCAGgcaga--cagaG-cAGCaCagGCaATTGTTttaat-c-CAGGA-Gcg

6 1048 GAcCaCCATTATGCGACTGTGcAGGACCTAAaACGAAAGTATTTAGGtAGTCCATATGtTAGTCCCTAT
11 1048 GAtGCTCATTATGCGACTGTGcAGGACCTAAaACGAAAGTATTTAGGcAGTCCATATGtTAGTCCCTAT
30 33 1104 GAgGATgATTtAaATGCTGTGtGtGcAcTAAaACGAAAGT TTGcCg
16 1093 aAncAACATagAGATGCaGTaCAGGTTCtAAaACGAAAGT AT TtGGTAGTCCa
31 1087 gAggAACATGcAGAgGcTGTGcAGGTTCtAAaACGAAAGT ATgtAGTAGTCCt
35 18 1139 cAcaAtgATGcAcAaGtGtTGcAtGTTtAAaACGAAAGT ttgcaggagcgacacaga
con gA-gatcAtT-agaggctgTgcagGttcTAAaACGAAAGTatttagg-agtcca--tgtga-tgcc-t
BE1-XAAaACGAAAGX-BE1
BE2-AGGACcXAAaACGAAAGXAXXAG-BE2
BE3-AGGXcXAAaACGAAAGXAXXGG-BE3
BE4-AXGXXXXAAaACGAAAGXAXXGAG-BE4

```

6 1116 AaCacTaTAGCgAgGCAGTgGAAAGTGAaATAAGTCCACGATTgACGCATTAAACTTACAAGAC  
 11 1116 AAGCAaTgTAGCTaATGCAGTaGAAAGTGAaATAAGTCCACGgTTaGACGCATTAAACTTACAAGAC  
 5  
 33 1151 ATGtccacaAagTGCtGCGAgGAGcTGTtGAtcGTgCTGcaAaacCCgtGtAgAacgtCTATtAAaTa  
 16 1146 cTTAGTGATATTAG TGGaTGTGTaGacaATAATATTAGTCCtAcgATTAAAGCTATATGTa  
 31 1141 tTAAGTGATATTAG TgTGTGTGTGATtATAATATTAGTCCACGCTTAAAGCTATATGTa  
 10  
 18 1198 aaAcagtcctATTAGgggagcggctggagGTGGATacagAgTtAGTCCACGGTTAcAGaaATATctt  
 con a-aca-tatAttagagggcagtgga-gtGtggtatgt--t-taaagtc--g--taaaagctAta-gta  
 15  
 6 1184 AGCCAAAAGGTAAAGCGACGGCTGTTTcAAAcCGGAAcTAACGGACAGTGGATATGGCTATTCT  
 11 1184 AGCCAAAAGGTAAAGCGACGGCTGTTTgAAACAcGGGAATTAACGGACAGTGGATATGGCTATTCT  
 20  
 33 1219 AaAtAAAgAAtGcAcAtacaGAAAAcGaaAAATAgTcAgcTAGAAGACAGCGGATATGGCAACT  
 16 1207 TAGAATAAaAAAGTAgAgCtGCAAAAGGgAgAtATTTCAAagcGAAGACAGCGGATATGGCAATAGT  
 31 1202 TAGAATAATaACAGTAAACAcGAAAGGAAAGACTcTTTGAAcTtcCAGACAGCGGATATGGCAATAGT  
 18 1266 TAAAtagTgggcagAAAAGGCAAAAAGGCGcCTgTTTAcAtatCAGATAGtGGcTATGGCgtTtCT  
 25  
 con -a-aaAaaag-g-Aaa--ag-aaaa-g-a-satatttgaaacta-caGAcAG-GGATATGGC-aT-CT  
 30  
 J33-tatggctattct  
 C87-ATACCGTTANGA  
 C88-ATACCGAYANGA  
 35  
 6 1252 GAAGTGGAGCTGgaacgggAAGC CAGGTAGAGAAACA TGGCG  
 11 1252 GAAGTGGAGCTG CAAGC CAGGTAGAGAAACA TGGCG  
 33 1287 GAAGTGGAACT CAGCAGAT GGTa CAACA GGTAG  
 16 1275 GAAGTGGAACT CAGCAGAT G tTACA GGTAG  
 31 1270 GAAGTGGAACT gCAGCAGAT G gTACA GGTAG  
 40  
 18 1334 GAAGTGGAAgC aaacCAGATTcaggtaacTACAAatggcgaaacatggcggaatgtatGTAG  
 con GAAGTGGAA-CTg----gca-caGataggtagag--ACA-----GtAg  
 gaaagtgaagctgghnnnnnchacagat-JJ3  
 45  
 CTTCACCT-C87  
 CTTCACCT-C88  
 50  
 55

```

6 1295          taCCGGAAAAATGG          GGGAGATGGTCAGGAAAAGGA
11 1289 A          cCCGGAAAAATGG          GGGAGATGGTCAGGAAAgGGA
33 1321 A          AagtcAAAAATGG          cgaCAC AaActtaaAtGacttaGA
16 1306 A          A          gggcGccatgagactgAAACACcAtgtagtctagtAtagtGg
33 1301 A          GGAG          CAAC          AAACA          AC
10 18 1396 tggcggcagttacGGAGGctatagaCAACGgggggcacagagggcAACA          AC
con      a-----ggag-----aacgcaaaatgg-----aga---gaaacacgagatggtcaggaagggga

6 1329 CACAGGaAGGGACATAGAGGG          GGaGGAACATAcAGAGGCGGAAGGccccacaaACAGtgtac
11 1323 CACAGGgAGGGACATAGAGGGTgagGgGGTGGAAcATagAGAGGCGGAAGCagtagacGACAGscccc
33 1358 atCtAGTGGGGTgGGGcAtGaTtcaGAAcTAAGcCTGtagacaaatGtAgaTagctGTGAAA
16 1349 tggAAGTGGGGGgtGgttcagTcagTAcAGTAGTGGaagtggGGGAGagGgTgtTAGTGAAGAGcAcA
20 33 1317 AttAAGT          tgtaATGTGTAGTG          ACGGGA          cAcATAGTGAACGagAgA
18 1445 A          gcagtgtagacggTacaAGTG          AC          aaTAgaatAtaGaaAat
con      a-caagtaggacagaga-ggt-aggga-g--agtgataga-cgg--gaagca---agtgAaaga-a--

6 1391 GgGAGCATGCAGgCACAgCAGGAATAT          TgGAATTgtTAAaATGAAaGATtTAC          GggCagCATT
11 1391 GaGAGCATGCAGACACatCAGGAATAT          TAGAATTAcTAAaATGTAAGGATATAC          GatCtACATT
30 33 1420          atgttACgttgCAGGAa AT          TAGTAATGTTCTAcAtagTAGTAATACAAaAGCAaAatAT
16 1417 cTatAtgcCaAAcACcacttacPA          ATATTTTAAATGtaCTAAAAACTAGTAATGCAAAgGCAGCAAT
31 1361 aTgAAaCtCCACAC          Gta ATATATTgcaAGTGTAAaAACTGCAATGgtTAAAGctcGCTAT
35 18 1487 gTAAAtCcCaCAAtgtaccataGcAaAtTAAaagAcTGTGTAAaAgtaAaCAATaaacAAGGaGCTAT
con      gtgaa--t-caa-c-ca-caggaAtAtattagaaatgtt-LAaaaa--ag-aTacaanaagcgc--aT

6 1455 AcTGGTAAGTTTAAaCAaATGCITTCGGGTGTCtTTTATaGATTAAATAGGCCATTAAaAGTGATA
40 11 1455 ACaTGGTAAGTTTAAAGAcTGCITTCGGGTGTCaTTTgTtGATTAAATAGGCCATTAAaAGTGATA
33 1478 ATTAtaTAAATTTAAAGAGgcCTATGGaaTaAGTTTtTgGAATTAGTAAGACCATTTAAaAGTGATA
16 1484 GTTAGcaAAATTTAAAGAGTTATAcGGGcTgAGTtTTtcaGAATTAGTAAGACCATTTAAaAGTGATA
45 31 1422 GTTAGgtAAATTTAAAGaATTTATATGGcGTAGTtTTtAgGAACtAaATTAGGCCATTtAAAGCAATA
18 1555 GTTAGcagTATTAAAGAcacATATGGgcTAtcaATTtAcAGattTAgTTAGSaactTTAAaAGTgATA
con      -ttaggtaaaTTAAAGA-t--TatGGgcT---tTTTatAGa-tTA-TAG-ccaTTTAAAGTgATA
JJ4-ttaggttagaccattttaaagtata

```

6 1523 aaACaACATGTTtaGATTGGGTGGTgCAGGgTTTGGTATACATCATAGCATATcAGAGgCATTTCAA  
 11 1523 gAAcACcATGTgCcGATTGGGTGGTtGCAGGATTtGGTATACATCATAGCATAGCAGaTGcATTTCaA  
 33 1546 AAACAAgCtGTaCcGATTGGGTATaScAGGATaTGGAAATTagTCCatcagTAGCAGAAAGTTTAAAA  
 16 1552 AAtCAACgTgTtgcGATTGGGTATtGCTGCATTtGGAcTTACACCcAgtaTAgCTGAcAGTaTAAAA  
 31 1490 AAAGCACAgtTACTgATTGGTGTgTACGTGcgtTTGGAGTtTACAggtACAgTtGCAGAAAGGATTAAAA  
 18 1623 AAACCAcGTGTACaGATTGGgtTAcAGCTataTTGGAGTaaAcccaACaTaGCAGAAAGGATTAAAA  
 con aaAc-AcaTGTacaGATTGG--t-tagC-ggaTtTGGaaT-satccta-aaTagCaGaggaTTtAAa  
 aaacaaCnTgtNcagattgg-JJ4  
 6 1591 AAaTTAAATtGAGCCATTAAgTTTATATGCACATATACAATGGCTaACAAATGCATGGGGAATGGTATt  
 11 1591 AAgtTAAATtGAGCCATTAAgTTTATATGCACATATACAATGGCTaACAAATGCATGGGGAATGGTATt  
 33 1614 gtATTAAATtAAACAgcATAGTtTgTATAcTcATtTACAATGTTTAACTtGcGataGAGAAATaaTaaT  
 20 1620 ACACtATtTaCAACAATATTGTTTATATTtAcAcATcCAAAAGTTTAGCATGTtCaTGGGGAAATGGTtGt  
 31 1558 ACcCTATTtCAACCATATTGTTTgTATTgCCATtTaCAAAAGTTTAGCATGTtCcTGGGGcATGGTTAT  
 18 1691 ACACtTaaTaCAGCCATTtTataTTATATgCCATaTcCAATGtCTAGAcTtGaaatTGGGGagTatTAAT  
 con aca-TAAATtcA-Coat-tagtTTaTATgcaCAtaTaCAATGt-Ta-catgtgcAtGgGaaTggTaaT  
 6 1659 gTTAGTATTAtTAAGaTTTAAAGTAAATAAaAGTAGaaGTACCGTtGCACGTACAcTtGcaACGcTAT  
 11 1659 ATTAGTATTAAATAGgTTTAAAGTAAATAAaAGCAGATGTACCGTtGCACGTACATtAGGTACGTTAT  
 30 33 1682 ATTATtGTTAAATtAGATTtAGgtGTAGCAAAAAACAGgtTaaCAGTaGCAAAAcTaTGaCTAAATTTAT  
 16 1688 GTTACtATTAGTaaGATaTAAATGTGGAAAAATAGAGaAACAAATTGAAAAATtGcTctTAAACtAT  
 31 1626 GTTaaTgCTtGTGAGATtTAAATGTGcAAAAATAGaaTAAcAAATTGAAAAATtTAAgaaAAATTTAT  
 35 18 1759 aTTAgccCTgtTGCctTAcAAATGTGgtAAgAgTAGAcTAACAgTTGctTAAAggtTTAAgtAcgTTgt  
 con -TTAgatfa-TaaGaTttAaatgt-gtAAaA-tAGa-taACAgTtGcaaaa-tatTaggtA-gtTaT  
 40 6 1727 TAAATATACCTGAAAAcCAaATGTTAAATaGACCCaCcaAAAAATCAAAAGTGGGTgGtgcACCGCCTGTAT  
 11 1727 TAAATATACCTGAAAAcCAcATGTTAAATaGAGCCtCtAAAAATCAAAAGTGGGTgGtgcAGAGCCCTGTAT  
 33 1750 TATcAATACCTGAAAcATGTATGtTaTAGAGCCaCcaAAATtACGgAGccAaAcAGtGGCATTGTAT  
 45 16 1756 TATGTgtTctCcaAtGTGTATGaTgATAGAGCCtCcaAAATtGCGTAGtACAGCAGCAGCATTATAT  
 31 1694 TgTGTaTATCTaCAaATTGTATGTtAAATCAGCCACCCaAAATtAcGTAGcACAGCTCAGCAGATTATAT  
 18 1827 TaccagTAcCTgaAAcTTGTATGTtAAATTCaCCACCAAAATtTgcGAGtAgTgtTCAGAGCacTATAT  
 50 con TatataTAcCTgaAAatgtATGtTAAT-gAgCCaCcaAAATtAcGagtagag--gcaGCa-T-TAT

6 1795 TGGTTTcGtACAGGtATaTCAAAATGccAGTACAGTTATAGGGGaaGCaCCaAATGGATAAcACGcca  
 11 1795 TGGTTTAGGACAGGcATtTCAAAATGcAAGTACAGTTATAGGGAGGgCCgCAATGGGATAAcGCGcca  
 5 33 1818 TGGTTTAGAAACAGGaATgTCAAAcATTAGTGATGTacAAGGtacAACCCGAAATGGATAgatAGAct  
 16 1824 TGGTATaaAACAGGtATaTCAAAATTATTAGTGAaGTgTATGGaCaACgCCAGAAATGGATAcAAAGACA  
 31 1762 TGGTAcAGAAcAGGAATgTCAAAcATTAGcGAtGTATATGGtGaaACACCAATGGATAgAAAGACA  
 10 18 1895 TGGTATAGAAACAGGAATaTCAAAATTATTAGTGaaGTaatggGaaGcACACCTGAGTGGATAcAAAGAct  
 con. TGGT-tagaaACAGGaATaTCAAAAttattAGTgaaGTaa-agg-gaaCaCCaGaaTGGATA-aaaGaCa  
 BE32-AXAXCAAAAXXAGXGAAGX-BE32 JJ6-tggataNaagaca  
 15 6 1863 aACaGTTATTGAACAcgGgTTGGCaGACAGTCAGTTTAAATTAAcGAaATGGTCAGTGGGcGTATG  
 11 1863 gACcGTTATTGAACATAGtTTGGCTGACAGTCAATTTAAATTAACTGAAATGGTCAGTGGGcATATG  
 33 1886 AACtGTTTTTCAACATAGcTTTAAATGATaTatATTGATtTAAgTGAaATGGTCAGTGGGcATATG  
 20 16 1892 AACAGTATTACAAcATAGTTTTAAATGATtgTACATTGGAATTATCaCAGATGGTACAAATGGGcTAcG  
 31 1830 AACAGTATTACAGCATAGTTTTAAATGAcAaACATTTGATTTGTCCcAAATGGTACAAATGGGcATATG  
 18 1963 tACtaTtTaTACaaCATGgaaTAgATGAtAgcAatTTTGGATTGTGcAGAAATGGTACAAATGGGcATtTG  
 con aACagTt-TacAaCAtaGttTt-ctGA-agttaaTTTGA-TTA-cagAAATGGTACA-TGGGcATatG  
 25 aacNgTataacacatagtttNgatgat-JJ6  
 6 1931 ATAATGAcATaTGCgAGGAGAGTGAaAttGCATTTGAATATGCACaaGgGGAGATTTTGAATCtAAAT  
 11 1931 ATAATGAtAttTGTGaaGAAAGTGAGAtAGCATTTGAATATGCACAGCGTGAGAcTTTGAATCtAAAT  
 30 33 1954 ATAACGAgTtAacgGACGATAGTGAcATTGCATAtAtATGcACAACTTGcAGAtTCaAAATAGTAAAT  
 16 1960 ATAATGAcATagTgACAGATAGTGAATTCGATATAAAATATGCACAAATTTGcGACACATAAATAGTAAAT  
 31 1898 ACaATGAGtTtATgCATGATAGTGAAATTGcCTATAAATATGCACAAATTAGCTcAGAGTgATAGTAAAT  
 35 18 2031 ATAATGAGcTgAcAGATGAAGGcGATATgGCaTtTgAAATATGcCcttATTAGCAGACAGCaAaAGcAAAT  
 con AtAAATGA-ataa--GA-GATAGTGAaATtGCAT-TgAAATATGCacatt-GCaGA--t-AtagTAAAT  
 40 6 1999 GCaGcGcATTTTTAAATAGcAAATATGCAGGCaAAATATGtgAAAGATTGTGCAAcATATGTGtAGACA  
 11 1999 GCaAgGGcCTTTTTAAATAGTAAATATGCAGGCaAAATATGtAAAGATTGTGCAATATGTGcAGACA  
 33 2022 GctgCTGCATTTTTAAAAAGTAACTCACAaGCAAAATATGTAAGGAcTGTGgATAAATGTGTAGACA  
 45 16 2028 GCaAGTGCCTTTcTAAAAAGTAATTcACAGGCaAAATATGTAAGGAcTGTGcAACAATGTGTAGACA  
 31 1966 GCAGcTGcATTTTTAAAAAGTAATTcGcAGGCaAAATATGtAAAGATTGTGgAACAAATGTGTAGACA  
 18 2099 GCAGcTGcCTTTTTAAAAAGCAATTGcCAaGCAAAAttTtAAAGATTGTGccACAATGTGCAaACA  
 50 con GCa-gtGC-TTTTAAAAAGTAAttcGcAGGCAAAA--tGTAaAGATGTGcaAaATGTGtAgACA

6 2067 TTATAAACATGCGAAGATGaggAGATGCTCTATaAAACAATGGATaAAacATAGGGGTCTAAAaTAG  
 11 2067 TTATAAACATGCGAAGATGaaAAGATGCTCTATaAAACAATGGATaAagATAGGGGTACTAAAgttG  
 33 2090 TTATAAAAAGCAGAAAAAagcTAAAAATGTCaATagGACAAATGGATAcAAagTAGATGTGAAAaAacAa  
 16 2096 TTATAAACGAGCAGAAAAaAaCAAAATGagTATGagtcCAATGGATAAAaATAGATGTGATaggGTag  
 31 2034 TTATAAACGAGCAGAAAAACGCAAAATGccATGgCAGATGGATtAAAagTAGATGTGACAAAGCTa  
 18 2167 TTATAggCGAGCccAAAAACGCAAAATGaatATGtcACAGTGGATagcAttTAGATGTtcaAAAAaTAG  
 con TTATAaac-aGCagAAAAa-ga-AaATGtctATyagaCAaTGGATaaaaataTAGetGTg-tAaa-tag  
 JJ11-tggataaaaatatagatgtNctaaaatag  
 6 2135 AagGcacAGGaAatTGGAAaCCAAATGCTaCAaTtCcTAcGACATCAAAATaTACAAATtATTCctTTT  
 11 2135 AcaGTCtAGGTAAcTGGAAgCCAAATGTCAGTtCTTAAGACATCAAAACATAGAATTTATTCcATT  
 33 2158 ATGATGGAGCAaATTGGAGaCCAAATaGTaCAGTtGTTAAAGATATCAAAACATcGAATTTAcagCATT  
 16 2164 ATGATGGAGGTGATTTGGAaGCAAAATGTCtATGCTTTTAAAGtATCAaaggtTAGAgtTTATGTCATT  
 31 2102 gTGacGAAGGTcactTGGAGGgacATAGTAAGTtTTTAAAGATATCAAAaATAGAAATTgTGTCATT  
 18 2235 atGaaGggGgaGATTCGAGagccaATAGTgcAAATTCcTgcGATAcCAACAAATAGAgCTTAaCAATT  
 25 con atgatggagG--AtTGGa--ccaAT-GTcagTtT-TaaGatAtCAaaa-aTGAaATTtct--CAATT  
 atgatggagaaattgga-JJ11 JJ12-cattt  
 6 2203 TTAActAAAtTtAAATTATGGCTGCACGGcACGCCaAAAAAAAcTGCATAGCCATaGTAGGCCctCC  
 11 2203 TTAAGcAAAcTaaAAATTATGGCTGCACGGaACGCCcAAAAAAATGCTATAGCCATTGTAGGgCCAGC  
 33 2226 TTAgtGTCATTtAAAaagTtTTTaaAAGGcATACCaaaaaaAagcGTATGcTAATTtTGGACcAGC  
 16 2232 TTAaCTGCATtAAAAaGaTtTTTgCaAGGcATACCTAAAAAAATtGcATATTAATATAGGTGCACG  
 31 2170 TTAAtCTGCATTAAAgctgtTTTTTAAAGGAGaTgCCaAAgAAAAAcGTATtTTAaTAcATGGTGCACc  
 18 2303 TTAGgaGcCTTAAAtatcATTTTAAAGGhaCCcCAaAAAAAAATGCTtTagTAtTtTgTGGAcAGc  
 con TTAa-tgcatTAAAtttaTtctT--AaGaa-gCCaAaAAAAa-TGtaTagtaAT-t-tGG-cCa-C  
 ttaagtgcattaaaattatttctgcaagNacNccNaaaaaana-JJ12  
 6 2271 aGATACtGGGAAaTCGTaCTTTTgATGAG TTTAATaAgcTTTtTaGCaGGcACAGTtATTAGTCaT  
 11 2271 tGAcACTGGGAAGTCGTCTTTTgATGAG TTTAATaAGcTTTtTgGGGGAAGcAGTtATTAGTTAT  
 33 2294 aAAACACGaaAGTCATattTTTGGaATGAG TTTAATaAGcTTTtTaaAGGGGTcTGTATATcATgT  
 16 2300 TAACACAGGTAAATCATtattTTGGcATGAG TTTAATGAAaTTTcTgCAAGGGGTcGTAATAgTtT  
 31 2238 TAATACAGGTAAATCATATTtTTGGaATGAgcCTTATtGAGCTTtTACaAGGATgTAATATCATAT  
 18 2371 aAATACAGAAATCATATTtTTGGaATGAGctTTAT acacTTTtTACaAGGAGcagTAATATCATAT  
 con -aATAcAGG-AAATCaTattTTTgGaATGAG-tTTAataaacTTtTAcagaGATc-gT-ATat--tAT

6 2338 GTAAATCCAGCAGCCATTTTGGTtGcAaCCgtTAgtAGATGctAAgGtaGCATTgtTAgATGATGC  
 11 2338 GTAAATCCtGCAGCCATTTtGGcctACAGCCAcTAAcGgATGCAAAAGtggCATTtTgGATGATGC  
 33 2361 GTAAATTCFAAAAGtCacTTTGGTTgCAGCCATTAtCAGATGCAAAATAGGAATgATAGATGATGt  
 16 2367 GTAAATTCFAAAAGcCATTTTGGTTACAAACCATTAGCAGATGCCAAAATAGGTATGTTAGATGATGC  
 31 2306 GcAAATTCFAAAAGTCATTTTGGTTACAAACCACtGtGctGATGCTAAAATAGGCATGTTAGATGATGC  
 18 2438 GtgAATTCcActAGTCATTTTGGTTggaACCgtTaaCaGATaCTAAggTgGcCATGTTAGATGATGC  
 con GtaAATTCcaaaAG-CATTTtGGcT-cAaCCatTAgcaGATgCtAAA-TaG-aaTgtTAgATGATGC  
 6 2406 AACACAGcCATGTTGGcATATATATGGATACATATATAGAAAtTgTTAGATGGTAATCCTATGAGtA  
 11 2406 cACACAACCATGTTGGACATATATGGATACATATATAGAAAAcctTATAGATGGTAATCCTATGAGcA  
 33 2429 AACgCcAatAaGTTGGACATATATAGATGATtACATGAGAAATGCGcTTAGATGGAAATgAAATTCcAa  
 16 2435 TACAgTGCcCTGTGGACtTACATAGATGAGcAaTtTAAGAAATGCGATTgGATGGAAATtTgATTTCTA  
 31 2374 TACAACCGCcAGTTGGcATtATATAGAcAaTAccTAcGAAATGGACtAGATGGCAAGcCTGTATCTA  
 18 2506 AACgACcAcGyGTTGGcacaTAcTtTcATAccTAtaTgaGAAATGCGcTAGATGGCAATCCaATaAgTA  
 con AACaccgccaatGTTGGacaTAtaTaGAta--tAtaTgaGAAAtgc-tTAcATGG-AAcc-at---tA  
 JJ15-gttggacatatatNgatacNtatatgagaaatgcggttagatgg-JJ15  
 6 2474 TtGAcAGAAAgCATAaAGCATTgACATTAATTAaATGTCCACcctTgCTaGtAGcGTCaAaCATAGAt  
 11 2474 TAGATAGAAAGCATAGAGcATTAAcATTAAATTAaTgTCCACCgCTaCTgGTTAcATCAAAATATAGAc  
 33 2497 TAGATGTgAAACATAGGcCAATTAGTgCAATTAaAATGTCCACCACtTgCTtTAcCTCAAAATAcAAAT  
 16 2503 TgGATGTAAAGCATAGAcCATTgGtAcAcTAAaATGcCTCCcATTATtAATtACATCTAAcATTtAAT  
 31 2442 TAgATGTAAAGCATAAAgCtTTAAATgCAGtTAAaATGTCTCCcTTATTgATTACATCTAAATATAAAT  
 18 2574 TtGATAgAAAGCAcAAAcCATTAAATcCAAcTAAaATGTCTCCcCAaTAcTAcTAAcCAcAAATATAcAT  
 con TAgaT--aaAgCATA-agCaTtTaatacaa-TaAAATGTC-CCAcTAcTa-TtAcATCaAAcAtaAaAT  
 6 2542 ATTActAAAGaAgAtAAATAtAaGtATTACATAcTAGAGTAcAACATTtACATTTCcAAATCCATT  
 11 2542 ATTAGCAAGAGAgGagAAATAcAAATATTACATAGTAGAGTtAcCAcATTtACATTTCcAAATCCATT  
 33 2565 GCAGGCACAGAcCTAGATGGCCCATATTACATAGTAGATTtAaCAGATTtTgaaTTTaaAAATCCATT  
 16 2571 GCtGGTACAGATTCTAGgtTGGCCcTATTTACATaATAGATtGGTGGTtTTTACATTTCcTAAATgagTT  
 31 2510 GCAGGTAAAGTgAcAGATGGCCATAcCTACATAGcAGAcTGGTGGTtTTTACATTTCcCAAAATCCATT  
 18 2642 CCAGcAaAGGATaAtAGATGGCCCATtTThgAaAGAGaATaacaGTaTTTgAaATTTCcAAATgCATT  
 con gc-ggtAaGAtgatAgaTggccaTAttTAcAtAgTAGA-TaacagtaTTTAcATTTCcAAATCCaATT



6 2610 CCTTTTGACAGAAATGGGAATGCAGTgTATGAACtGTCAaATcAAAACTGGAAATGTTTTTTGAAA  
 11 2610 CCCcTTTGACAGAAATGGGAATGCAGTATATGAACtATCAGATGCAAACTGGAAATGTTTTTTGAAA  
 5 33 2633 CCCATTTCATGAAATGGtAAcCCAGTGTATGcAaTAAATGATGAAAACTGGAAATCCTTTTCTCAA  
 16 2639 TCCATTTCAGAAACGGAAATCCAGTGTATGagcTtAAATGATAAAGCTGGAAATCCTTTTCTCAA  
 13 2578 TCCATTTCAGAAACGGAAATCCAGTATATGAATTAAGTGATGAAAACTGGAAATCCTTTTCTCAA  
 10 18 2710 TCCATTTCATGAAATGGcAATCCAGTATATGAaTAAATGACaAAAAATGGAAATGTTTTTgAAA  
 con -CCaTTTGacaaAAATGG-AAtcCAGT-TATGaacTaaatgAtaaAAcTGGAAAT-TTtT---AA  
 6 2678 GACTGTGCTCAAGCCTAGACATTcAGGATtctGAGGA CGAGGAA GATGGAAGCAATAGCCAA  
 15 11 2678 GACTGTGCTCCAGCCTAGACATTgAGGATTCAGAGGA CGAGGAA GATGGAAGCAATAGCCAA  
 33 2701 GGACGTGGTgCAaATTAGATTTaataGAGGAAGAGGA CAAGGAAACcATGGAGGcAAATcAGc  
 20 16 2707 GGACGTGGTgcAGATTAAgTTTGACAGAGGAGAGGA CAAGGAAACcATGGAGACTCTTgcCA  
 31 2646 GGACGTGGTgCAGATTAAATTTGCACGAGGAAGAGGA CAAGGAAACcATGGAGACTCTTtcCA  
 18 2778 GGACaTGGTcCAGATTAgATTTCACAGAGGAAGGAagatgcAGAcAcCGAAGCAaACCTTTCggA  
 con GgacgTgTccAgatTAgattTgcscGAGgaGAGGA---c-agGaaacgATGGA--ca-T-tcc-a  
 6 2740 GCGTTTAGATGCGTGCCAGGAaCAGTGTGTAGAaCTTTATGAAGAAACAGTAcTGAacTACACAaAAC  
 11 2740 GCGTTTAGATGCGTGCCAGGAtCAGTGTGTAGAaCTTTATGAAGAAACAGTAcTATATACACAaAAC  
 30 33 2766 ACCTTTAAATGcagtgCAGGAaAAAAIACtAGAcCTTTAcGAAGcTGATAaaACTGATTAcctcAC  
 16 2772 ACCTTTAAATGTGTGTcAGGACAAAAATAcACATTTATGAaATGATAGTACAGAGcTACGTGAac  
 31 2711 ACCTTTAAATGTGTGTcAGGACAAAAATAcGAACATTATGAaATGATAGTAAAGcATctGTGAtC  
 35 18 2846 ACCTTTAAgTtgctgCAGGACAAAAATcATAGAcAcTATGAaATGAcAGTAAAgacTagacagcc  
 con aCGTTTAAaTgcgtg-CAGGAcAaaTatTAgaac-tTATGAA-atgA-AgtA-gaccTagacaaAC  
 6 2808 AtgTatTGcATTGGAATGCAATGcAGAcATGAAAGTGATTATtATtATtAAAGCAAAACAAATGGCCtA  
 40 11 2808 AcATATgCATTGGAATGCAATGcAGAcATGAAAGTGATTATtATtATtAAAGCAAAACAAATGGCCtG  
 33 2834 AaATGGAaCATTGGAACtGTATACGCAAGcAGTGTGCTTTATTgTATAcAGCCAAACAAATGGGATT  
 45 16 2840 ATATAGACTATTGGAACAcATGCGCTcCAATGTGCTATtTatTAcAAgCCAGAGAAATGGGATT  
 31 2779 ATATAGACTATTGGAACAcATtCGAcTtGAATGTGtATtAaTgTATAAGCAGAGAAATGGGAATA  
 18 2914 AaATAGAcgATTGGcAACTaATaCGttggGAaATGcAATATtTtTgcAGCAAGGGAaAcTGGcATA  
 50 con AtaTgag-ATTGGAAC-cATAcGactgCAa-gTG-atTatt-tataaaGCAa-a-AAatgGG--Ta

6 2876 AGGCACATAGGaaTgCAAGTACTGCCACCATTAAAgGTGtccGaaGCaAAAGGACATATAGCcATTGA  
 11 2876 AGGCACATcGGgtTtCAAGTAGTcCCACCATTAAcTGTGTcAGAGACTcAAAGGACATATAGTcATTGA  
 33 2902 tCACTtTatGCCACCAGtGGTGGCtcttCTTgttagcATCAAGACcAAAGcATTTCAGCAATTGA  
 16 2908 AAACATATTAACCAcCAAGtGGTGGCCAAcCTGgCTGTATCAAGaaTAAAGCATTCACAGCAATTGA  
 31 2847 CAcATATTAACCAcCAGGTGGTGGCAGGtTGTtCaGTATCAAAAGccAAAGCcTTACAAAGCATATTGA  
 18 2982 CAgAcATTAACCAcCAGGTGGTGGCAGCCTatacaATtCAAAaagtAAAGCacataAAAGCATATTGA  
 con aaccataTaa--ccacCA-GTgGTGCCa-Cattgac-gtaTCaaAgactAAAGc--t-AAGcATTGA  
 J118--tcaaaagactaaagcacataaagcNattga  
 6 2944 AATGCAAAATGCATTTAGAATCattTattAAggACTgAGTATaGtTcTGGAAcCGtTGGACATTACaaGaaA  
 11 2944 AATGCAAAATGCATTTAGAATCctTTAgcAAAACTCAGTATgTgTGGAAcCCTGGACATTACAgGacA  
 33 2970 ACTaCAAAATGgCaTTAGAgACATTAAgTAAATCACAGTATAGTACaagcCAaTGGACATTgCAACaaA  
 16 2976 ACTgCAAcTAACTGTAGAAACAACTAtATAACTCACAATATAGTATGAAaAgTGGACATTaCAAgAcg  
 31 2915 ACTaCAAAATGAtGTtGgAAACAtTAAATAAcACTgAAATACAAAAATGAGGAcTGGACAAcTGCAGcAAa  
 18 3050 ACTgCAAAATGgcccTaaAAggccTtgcaCAaAgTcgATACAAAaccGAGGATGGACaCTGCAGAcA  
 con ActGCAATg-c-tTAGAaAcATa---aaactca-TAtagta--gaaca-TGGACATt-CAAgA-a  
 actgcaaatgg-JJ18  
 6 3012 CAAGTTATGAAATGTGGCaAACACCACC tAAACGcGTGTTTAAAAAacgGGGcAAaACTGTAGaAGT  
 11 3012 CcAGTTATGAAATGTGGCTAACACCACC CAAACGgTGCtTTAAAAAaCAGGGAaACTGTgAGgGT  
 33 3038 CAAGCTTAgAGGTGTGGCTtTgTgAACACC AaaATGTTTTAAAAAaCAAGaAAaAGTaaCTGT  
 16 3044 ttAGCCTTGAAGGTATTTAACTGCACCAC AGGATGTaTAAAAAaCATGGATATACAGTgGaaGT  
 31 2983 CAAGCTTGAACGTGATTTAACTGCACCTAC AGGgTGTTTTAAAAAaCATGGATATACTGTAGAgGT  
 18 3118 CAAGGagGAACtATggaatACaGaACCTACTcactgCTTTAAAAAa gGTGGcCAaACAGTCAAGT  
 con caaG-t-tGAA-TgTggctaac-gcACCaacaa-g-tgttt-AAAAAacatGGA-A-AC-GTAGaaGT  
 6 3079 TAAATTTGA TGGCTGTGcAaCAATacAATGGATATGTGGTATGGACAgTgTgTatGTGCAGG  
 11 3079 aAAATTTGA TGGCTGTGAAGaCAATgtaATGGAGTATGTGGTATGGACAcATATATAcTGCAGG  
 33 3105 GCATATGA caatGACAAAaAaATACAAATGGATTATACAAACTGGggtgAAATATATATTaTg  
 16 3111 GCAGTTTGTATGG aGATATgCATGCAATGCATTATACAAACTGGAcAcATATATATTATTCTG  
 31 3050 GCAATTTGTATGG tGATGTAcACAaAcCATGCATTATActAACTGGaaATTTATATAcTATGTA  
 18 3185 atAtTTGTATGGcaacaAaGacaAttgtAtgAcctATgtagCatggygaAGtgTgtatTataTgacTg  
 con gCAATtGTATggcaacgatgaasacaatacaAtggAttat--caaaactggacagataTaTAtg--tg

6 3144 ACAATGACaCetGGGTAAGGTGcaTAGTatgGTAGATGCTAAGGctATATATTACATCTGGACAA  
 5 11 3144 ACACAGACtCATGGGTAAGGtaACTAGTtccGTAGATGCCAAGGGcATATATTACATCTGGACAA  
 33 3170 AgCAAGAtaCATGtACTATGGTtACaGGGaaAGTAGATtATaAGTATgTATTATATACATaAGtGt  
 16 3176 AAGAAAGcatCagtaACTCTGCTGtaGAGGctCAAGTtCACTATtAaGGTtTATATTATCTtCATGAAGGA  
 10 31 3115 TAGAtGGccatTGTACTGTtGTgGaaGGGcAAGTtAAATtGTAAAGGGcATtTATTATGTACATGAAGGA  
 18 3253 atGcaGGaacATGggcaaaaccGctacctgtGTAAgTcacAgGGGatTgTATTATGTAAAgGAAGGg  
 con a-gaaGacacatgg-cta-ggt-g-t-gt--aAGTagattataagGGtaTaTATTAt--tacaagaagg  
 15 6 3212 TTTAAACATATTTATGTAAAcTTTgtaaaAGAGGCagAAAGATATGGgacACCAaaCATTTGGGAAGT  
 11 3212 TTTAAACATATTTATGTAAATTTTAAAtAAAGAGGCacAAAGATATGGTAGtACCAAtCATTTGGGAAGT  
 33 3238 gaAAAggtATATTTTaaAtATTTTAAAGAGGATGtCCcAAAGTAtcTAAAAcacaAAATgTGGGAAGT  
 20 16 3244 atAcgAACATATTTTGTgcaGtTTAAAGAtGATGCAGAAAAATATaGTAAAAaTAAAGTATGGGAAGT  
 31 3183 cATatAACATATTTTGTAAAtTTTACAGaAGAGCAaAAAAATATGGGActgTAAJAAaATGGGAAGT  
 18 3321 tAcAcacAGtTtTTaTatAgAaTTTAAaAgTGAatgtgAAAAATATGGGAcacaggtAcgtGGGAAGT  
 25 con t-taaaacaTaTT-TgttaaAtTTTaaa--aaAGggcagaaaa-TATgq--Aa-ac-aaaaa-TGGGAAGT  
 6 3280 ATCTTATGGCAGCAGAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTacACaGAAGTAT  
 11 3280 ATGTTATGGCAGCAGAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTgtaCgGAAGTAT  
 30 33 3306 ACATgtGGGTGGTCAGGTAAAT tgTTTGTCTAC gtCTaTATCTAGCA ACCA AaTAT  
 16 3312 cATGCGGGTGGTCAGGTAAAT atTATGCTCTACATCTGTgTTTAGCAGCA ACGA AgTAT  
 31 3251 gCATGCGGGTGGTCAGGTAAAT TTTCTCTgAATCTGTaTTTAGCAG TGACGA AaTAT  
 36 18 3389 aCATtttGGGaaTaaGTAAATTGattgtTaaTgactctatgtgcagTACAG TGACGAcacggtAT  
 con acaT---GGT-gt-agGTAAAtg-at-tt-Tcctgacac-tct-t--c-AGcactgac-aagaagTAT  
 BE21--CGGTAT  
 40 6 3342 CCAITtCCTGAA tCTACTACATACACCCCGGCACAGACC tCcACCCt tGTGTCTCtCaagC AC  
 11 3342 CCAITtCCTGAA CTTACTACATACACCCCGGCACAGACCacCCcCTTacaGtGTCCgctGc AC  
 33 3362 CCACTTaCTGAAACTgCTgACATACA gACAGAC AAC gaTaaCCGACC AC  
 16 3371 CCTCTcCTGAAAtTATTAggcagCA cttgGCC AACcAcCCGcCGACCCcATAC  
 45 31 3310 CCT tGCTggaATTGTtAcAaAGGTACcaacaGCC AACaACACCACCAcAtGaATTC  
 18 3454 CC gCTaCTCaGctTGTtAAaAGCTAC AgcACACCcCCTCacCGtATTC  
 50 con CCact-cTgaaa---ttgacatacAcccagcagcagacc--c--caacaaac-cctcc-CaaCC-ataC  
 CC---GCXACXGAGCXG-BE21

6 3403 CA AGGAAGACGcagTGCaaaCGCGCCCTAGGAAaCG  
 11 3406 CAC GGAAGACGCGcGTtcgCGCGCCCTAGGAAaCG  
 33 3410 CACAAGC agcgggcAAACGacGACGAC cTgCAGacACCA  
 16 3426 CA AAGCCgtCGCCTTGGGCACC GAAG AAACacagGACGAC TATcCAGCGACCAAGA  
 31 3368 CA AAACctgCGCCTTGGGCACCcaotGAAGgtgTgCGGCGggcGACGACGCTCtactaAGCGACCAAGA  
 18 3503 CA gcACCgtgtCCgTGGGCACC GcaAAgacctACGCGC caGACGTC  
 BE10-GGCGXGXCGCGCCGCCXAG-BE10  
 con CACAaacctgtgccttgggcacc-g-gaaggcgtacgaagac-gacgacgtcc-cc-ag--accaaca  
 6 3439 AGCACGaggagtccacaCAgTCCcCtTgCAACgCctTGTGTGTGGCCcACATtGtGacCCGTGGACAGTg  
 11 3442 AGCAGC tggACCGTcCaTaaCAACaCCcTGTGTGTGGCCcACATcAGatCCGTGGACAGTA  
 33 3449 CAGACACCGCCcAgCCCT tacaAgtGTtTcTGCA gaCccgCCTTGGCAaTA  
 16 3481 tCAGAGCCAGACACCG GAaACCCCTgCCACaCCaCTaAGTTGTTCGacagaGACTcAGTGGACAGTg  
 31 3435 aCAGAGCCAGAGac aGAaACcCCcACcCCaAcAAGTTGTTCGAGcGCGACTcCGTGGACAGTg  
 18 3548 GgtgctAcACgaCCTggACaCtgtggaCTcGcgGaGaaGcAGcattGTGGAC cTg  
 con acaaaagccaga--ccgc-aaCCcct-c-acaccatgt-ttgggtgcacagcggtccgTGGACagTg  
 6 3507 gAAacCACACcCTCaTCACTaAC AATcAGcAcAGACCCAA GAcGG AACACaG  
 11 3504 CAATCaACAACATCGTCACTgAC AATtACAACAAGACCCAAA GAAG AACCACTG  
 33 3506 gAACAgcaCgtACTGCAACTAACTGC aCAAAcAAGCAgCGGA cTgtGTGT AgTTc  
 16 3548 cTCCaATcCtCACTGcAttTAACAGCT CACACAAAggACGGA tTaaCTGT AaTag  
 31 3502 TCAACTgtgggTtATcAGTGCAGCT gcatgCACAAAcAAACAA GGgCTGTcAGTtGTcc  
 18 3604 TCAAC ccactTcTcGTGTCAGCTacacctacagggcaacAAcAAaagacCGaaactCTGTtGTg  
 con c--aac-c--cactgc-actaaCagctaact-c-aacaagcccca-Aa-----gggtgtcaaca-t-g  
 6 3562 TaACAGTtCAGCTACGCCCTATAGTGCaaTtCAAGGTGAaTCCAATTGTTTAAaGTGTTTAAgATATA  
 11 3559 TcACAGTGCAGCTACGCCCTATAGTGCaaTtCAAGGTGAaTCCAATTGTTTAAaGTGTTTAAgATATA  
 33 3561 TAACgtTGCA CCTATAGTGCATTAAaAGGTGAaTCAaAATAGTTTAAaGTGTTTAAgATATA  
 16 3603 TAACACTACA CCCATAGTACATTTAAaAGGTGAaTCAaAATAGTTTAAaGTGTTTAAgATATA  
 31 3563 TgcACTACA CCTATAATACaTAAaAGGTGAaTCAaAATATaTTAAaGTGTTTAAgATATA  
 18 3668 TaacACTAC gCCTATAATACaTAAaAGGTGAaTCAaAATAGTTTAAaGTGTTTAAgATATA  
 con TaacacTaCagctacgCCtATAgT-CattTaaAAGGTGatccaAatagTTTAAaGTGTTTAAgATATA  
 JJ20-catttaaaaggtgancNaatagtttaaaagtgttaagatata

6 3630 GgCTaAATGACAgAcAgacATTtTaTTGAtTTaAtATCATCAACGTGGCAcTGGGCTCtCaaAG  
 11 3627 GaCTgAATGACAAATATAAACATTtTGTtTGAAtTAgcATctTCAACGTGGCAcTGGGCTCacCtgAG  
 5  
 33 3621 cAGATTaAAAGctTcATAAgAgTGTATAGTTCtaTGTCaTCCacCtGGCAATTGGACAGtgacAAc  
 16 3663 TAGATTtAAAAAgcATTgtacATTGTATATgTCAGTGTCTCTACATGGCAATTGGACAGgacATAA  
 31 3623 TAGgCTGtCAAAATATAaacAATTGTATgAcAAAGTGTATCTACATGGCAATTGGACAGtGacagAT  
 10  
 18 3728 cAGAttGgAAAAcATTAgcgAccacTATagAgAtaTATCATCCacCtGGCAATTGGACA ggtgc  
 con g-egattt-aaaaa-Ata-aca-ttgTaT-a-t-a-t-TCaTC-AC-TGGCAcTGGaC-tg-cc-aa-  
 gcagatt-JJ20  
 15  
 6 3698 GCACCACATAAA cATGCCATTGTAACTgtTAACAT ATgAtAGTGAGGAACAAgGCAcAgTTTT  
 11 3695 GCACCACATAAAA ATGCAATTGTAAcAtTAACAT ATAgcAGTGAGGAACAAcGcAgCAaTTTT  
 33 3688 aaaaAtgtTAAAA ATGCAATTGTAACTgtTAACATTGtaAcTGAACAGCAACAAc AaatgtTTTT  
 20  
 16 3730 CtAAAAcATAAAA gTGCAAATTGTtACacTtACATATGatAGTGAATgGCAAC GtGACcAaTTTT  
 31 3690 GGAaAACATAAAA TGCTATTGTaaCctTAACATATataAGTAcATCACA AGAGACgAtTTTT  
 18 3791 aGgcAAtgAAAAaGgaATacTgACtGTAAcATAccatAGTgAaCACA AGAacaaAtTTTT  
 25  
 con g-a-aaCctAAAAsatGCAATTgtTAACTgtTAACATatgatagt-aa-aggCAcraeg--aaCaTTTT  
 6 3762 TagAtgtTGTAAAAATACCCcCtACCATTAgcCA cAAcTGGGAATTtATGTCAcTgCAcctATTGTA  
 11 3759 TaaAcAgTGTAAAAATACCAcCCcACCATTAGcCAT AAggTGGGgTTtATGTCAcTtCATTTATTGTA  
 30  
 33 3752 TaggtAccGTAAAAATACCAcC tACTgTgcAAAT AAg TACTGGGATTAT  
 16 3794 TgtcTcaagTtAAAAATACCA AaaACTAttaCAGT GTC TACTGGGATTAT  
 31 3754 TAAATACTGTtAAAAATACC tAAcACagTatCAGT GTCaaccagatatatgactATTTA  
 35  
 18 3856 TAAATACTGT  
 con TaaatactGTAAAAataccacaaGaca-tagcaat-aaggctgg-tttatgt-actg-atttatgtga  
 6 3829 AtttgtatatatgttaaAtgtgTaaATATATGTATgtGTGTaatacaActgtTACATGTATGGAAgTgG  
 40  
 11 3826 A ccATTACaccTgttATATATG TATATGTGTA CATAAcATACgGTGTATGGAGtGtAG  
 33 3801 G aCATTA TAAgTGTA CATACAAgCcaATATG  
 16 3843 G  
 45  
 31 3812 tCt  
 gCc  
 18 3866 tgcaattccagtagtgtgacaaatattgggtgggataCa  
 con a-----g-catta---t---atatggttatatgtgta---cctaacaacatgtatggaagctcg

```

6 3897 TGCCTGTACAAATgCTGCAGgAACCAACcAgcACATTcATAcT      GCCTGTTaTaAATTGCAT
11 3881 TGCCTGTACAAATgCTGCAGcAACAACTAcAACATTGATATT      GCCTGTTGTTATTGCAT
5
33 3833      TGCTGC TAAATgTATAcAACcATGATAITTTgTTTTTG taTTaTGTITTTaTaATT
16 3847 aTATGA cAAATcTtgatactgcATccAcAaCATTAcTgGcgTGCCTgTTTG CTTTGGCTTT GAGTG
10
31 3815 TAATGATtgAActaAataattTcTAcagtaAgCATT gTGCtaTGCTTTTTTG CTTTGGCTTTGAGTG
18 3904 TgAcAAIgtAAatacAtatgcTgTAgTaccAatATgttatCacTtaTTTTttatTTTGGCTTTTGTGT
con      tgac-atacaa-ttgctgc-tgaacaaccca-ca--tt-ata-TgctttttggccTtt-cTtttggttt
021-CTGCAGGAACAACCAGCACATTCTACT      GCCTGTTATAAATTGCAT
15
6 3957 TTGctGTATGTTtTtgTTAGcATcaTACTTAItgPATggATARCTGAgtTTaTcTgTAcAACACTGTG
11 3941 TTGcaGTATGTAATcTTAGcATtTgTACTTATaTaTATTAAATAcCTGAtTTTgTaGTAATACATCTGTG
20
33 3886 gT      TTTtaTgctTATcCtTATTATtTACGTCCctTTAAAcTTTCcaTTTCTACcTATgCtTggTTG
16 3911 CT      TTT GTGTGTCTgCcTATTAAATACGTCCcgCTgctTtTGTCGTGTCTACATAcCaCATCATTa
31 3880 CTactaTTT GTGTGTCT      tgcATACGTCCaCTgTgCTGTGTGTGcgTATAtgCaaCAcTA
18 3971
25
con      -tgctgtttg-tgtgt-tgcat--ta-taggtccatt-ataattttct-tttctgtatatacatc--tg
021-TTGTTGTATGTTTGTGTAGCATCATACTTATTGTATGGATATCTGAGTTTATTGTGTACACATCTGTG-021
30
6 4025 CTAGTACTAACACTgCTTTTATATTTsCtaTTGTCGCTgCTATTAAcAACCCcCTT GCAATTTTTC
11 4009 CTGGTACTAACACTTCTTTTATTTTTCCTTTTGGGCTtTATTAACAAACCCCTT GCAATTTCTTT
33 3950 CTGGT gTTGGTATTgCTgcTtTggGtgTTTGGG      gATCtCCTTTaaAAATT TTTT
16 3974 ATAAAT ATTGGTATT      acTaTTgTGGaTAACAGcAGcCTCTgCgTTTaG
35
31 3943 cTATt ATTaaTtgT      gaTtTTaTGGGgTtAttGCAaCCTCTcCaTTaG
18 3971
con      ctagtac-tt-attttttttatttggcttttggcctttatgaa-aac-cc-ttc-caattttt--
40
021-CTAGTACTAACACTGCTTTTATATTACTATTGCGGCTGCTATTAAcAACCCcCTT GCAATTTTTC-021
45
50
55

```

6 4092 TACTAACTCTactTGTGTGtTAcTgTCCcGCaTtGTATATACactacTatATTGT taccaca  
 11 4076 TACTAAcAcTtTgtTGTGTGCTATTTCCCTGcTTTATATACACatataTACATTGTgcaacscgcAacaA  
 5 33 4005 T tTGTCTASTT GTTgTtTTTATAT tTA CCAaTGAgtTgTATtAAttt  
 16 4021 gTGTtTTtaTTGTATATatTATT GTT TATA TA CCATTATTTTAAATcATATC  
 31 3990 tTGTtT TGTATATAT gTgT GTT TATA TAcattcCATTATtTgTAAATcATATC  
 10 18 3971  
 con tactaacactgtatat-tgctattttctgttc-ttttataactata-tccattgtttttaat-cata-  
 021-TACTAACTCTACTTGTGTGTACTGTCCCGCATGTATATACACTACTATATTGT TACCACA-021  
 6 4154 aGcgaTGATGCTAAcATGTCaATTtAAATGATGGaGAT ACcTGGcTGGgTtGTGTGTtTatgTG  
 11 4144 TAAATGgTGATGCTAAcCTGTGCACTtAAATGATGGcGAT ACATGGtTGTtTcTGTGTGTtTactTG  
 33 4051 TCATGCACAGcaTatgacacacAaagAgTAATGTATAT ACATGtataTATtTGTtGTATATatgTG  
 16 4075 ACATGCCAC gctTTTTAattacaTAATGTATATGACATAaTgtaATTGTT ACATATA  
 20 31 4044 ACATGCA tctTTTTTAA GTCAACAgTaaCTTTT AC  
 18 3971 AtgcattgtatgtgtgctGcAtgtcccgCTTTTgccAtctgtctgta  
 con -cstgcacatg-taac-t-t-t-aattaaataatggagatgacatggttg-tTtt-tg-t-t-tatgtg  
 021-CAGCAATGATGCTAACATGTCAATTtAAATGATGGaGAT ACCTGGCTGGGTtTGTGTGTtTATGTG-021  
 25 6 4220 CcTTTaTTCATAGggaTgtTgGGgTtATtATt gaTgCacTatAGaGCTGTACaAGGGgaTaAAc  
 11 4210 CATtTgtTGTAGcTGTactTGGATtGTtTtTt actaCATTAcAGgGCTGTACATGGTactgAAAA  
 33 4117 CA caTGGTgTGTtTtTAAcATTGtTtTt gTATtTtT AgtTtTtTtTtTtTGTa  
 30 16 4132 atTgTGTATACCaTaacttactaTtTtccTtTtTATtTtTcTaTataTtTtTtTtTtTGT  
 31 4081 TtGTGTATAC TgTtTgtTgTatTgGTattggtTaTggtTaTggt  
 18 4018 tgtgtgCtATgCAtgggtattggtatttgggtatatTgtggTaataaacGTccccctgccacagcaTtc  
 35 con catttt-tt-gtg-a-t-ctag--tt-ct-ct-t-ct-ttt-ctt--a--q-t-t-tttttt-ctt-t-  
 021-CCTTTATTTGTAGGATGTTGGGGTtATTATT CATGCATAATAGACGTGTACAGGGGATAAAC-021  
 6 4283 AcACcAAATGTaagAAGTGTAA CAAAC aCAactgTAATGatGATTATGCTAACTATGcattATacT  
 40 11 4273 AAaCTAAATGTgctAAAGTGTAAATCAAAcGcCAaactacTactGTgGATTATGCTATGCTaacATggt  
 33 4171 ttACTAA TAAAT AccTTTATatTtttagcaGTGTAT  
 16 4196 TtTgtTtTgtTtTgTTTTTTA  
 31 4124 tattggTATtggTaTaaTaaactTTTtTACtTtTTTTT  
 45 18 4086 acagtaTATgtaTtTgtTtTtTtTtTgTgCcaTgTtTAcTatgCataacatgcatattgtctttaca  
 con a-actaaactgattaaagtgaatt-t-cc-t-t-ttT-atgttgattaaagtgaattg--tatact  
 021-ACACCAAATGTGAAGAAGTGTAA CAAAC ACAaCTGTaATGATGATTATGCTAACTATGCTATTACT-021

```

6 4348 acTgATGGcGATTAT aTaTATATGAACtAGAGCTAAACCGTTTTTTATAtttgttaacagGTGTatGc
11 4341 gaTaATGGaGATTATG TgTAcATGAACtAGAGTAAACC TTTTTTATAcagtggtgGTGTacGt
5
33 4206 tATCTG
16 4214 ataaactgTTATTA
33 4164 TTATTA
10
18 4154 gtaattgtataggtgttttatacagtgattgtacattgtatatatttgttttataccottttaTgCTt
con g-taattggagattatg---tatacatgaa-tagagtaaacc-ttttttatatt-ttaat-gt-tatt-
021-ACTGATGGTGATTAT ATATATATGAATTAGAGTAAACCGTTTTTTATATTGTAAACAGTGATGC-021
15
6 4413 TttgTATAcATggcacAtagTAGGGccCGacGACGcAagCGTGCGTCAGCtACACAGtATATACAAA
11 4405 TagtTATA tATaatgAAACtTAGGGCCGcAGCtAAaCGTGCGTCAGCtACACAACTATATACAAA
33 4213 AGACACAAACGaTCTaCAAGgCGCA AgCGTGCGTCtGCaACACAACTATaC AAA
20
16 4228 CtttaacaATCGGACACAAAGGvTCTgCAAAAGCCACaAAACGTGCATcGcTACcCAACTtTATaAAAA
31 4170 C cATGCGGtcCAAAAGGcTCTaCAAAAGCCACTAAACGTGCGtCTcGTACaCAATtTATaC AAA
18 4222 tttgtattTttGtaattAAAGtattggtAtccCaCgyTgcccgaacgacgcaaacggygctTcggttaactg
con tttgttatat-aga-acaAaact-c-gcaagacgc-gtaaacgtgc-tc-gctacacaaactatatcaaa
25 021-TTTGTATACCACTGGCCACATAGTAGGGCCCGACGCAAGCCTGCCTCAGCTACACAGCTATATCAAA-021
6 4481 CATGcAAactCACTGGAACATGcCCCCCAGATGTAAATTCCTAAAgGTgGAgCacAacACaATTGCAGAT
11 4472 CATGcAAGGCCACTGGtACATGtCCCCCAGATGTAAATTCCTAAAGTtGAACtActTACtATTGCAGAT
30
33 4268 CATGcAAGGCCACAGGcAcCGcCCACCCGATGTTATTCCTAAAGTgGAAGGAgtTACcATaGCAGAT
16 4296 CATGCAAAcagcCAGGTACaTGCCACcTGaCtTATACCTAAAgTtGAAGGcAAaAcTATTCcTGAA
31 4233 CATGCAAAAgcAGCAGGTACTtGTCCaLcAGACgTTATACCTAAaAaTaGACaTActACaATTTCcAGAC
35
18 4290 actTatAtaaAaCAtGTAAacaaatCtgytacatgTccACCTgAtgTtGtLcCtAaggtgagGgcacC
con caTgcAAagcCaCagGTActgttcCaccagatgttat-CCTaAgTtGascataAaaccattGcagat
40 021-CATGTAAACTCACTGGAAACATGCCCCAGATGTAAATTCCTAAAGGTGGAGCAACACCACTTCGAGAT-021
45
50
55

```



6 4549 CAAATATTAAAAATGGGGAAGCTTgGGGGTgTTTTTTGGaGGGTgGGTATaGGcAcGgGtCCGcGAc  
 5 11 4540 CAAATATTAAAAATGGGGAAGCTTAGGGGTgTTTTTTGGGGGTTAGGGTATTGGcACAGGGgCTGGTAg  
 33 4336 CAAATtCtAAATATGGcAGTTAGGGGTTTTTTTGGTGGcTTAGGGTATTGGcACAGGGCTGGTtC  
 16 4364 CAAATATTAcAATATGGAAGTATGGGTGATTTTTTTGGTGGGTAGGAATTGGACAGGGTCTgGGTAC  
 10 31 4301 CAAATATTaaggTtAGGTAGTATGGGTCTTTTTTTGGGGGGTgGGTATTGGGTtCgGGCTtCGGTAC  
 18 4358 acgtTAGcAgataAaatattgcaatGgtcaagccTTGGTataT7TtgggTGGacttGGCataGGTAC  
 con caaaTattaaaaatgtgaagttt-gGggttttttTTGGTgggTTaggtattGG-acagGctctcGgtac  
 021-CAAAATATTAAAAATGGGGAAGTTTGGGGGTGTTTTTTGGAGGGTTGGGTATAGGCACGGGTTCCGGCAC-021  
 15 6 4617 TGGgGGTCTGtaCtGGcTATgTtCCCTTacaAAcTCTGcAAaACCTcCTATTACTaGtGGGcCtatgG  
 11 4608 TGGcGGTCTGtGcAGGTATATACCTTgGGAAGcTCTCCcAagCCTGCTATTACTGgGGGcCagcAG  
 33 4404 AGGTGGAAgGactGGcTATgTACCTaTtGGtActgacCCaCCtACAGGTgCAAtccCctTGcAgCCTa  
 20 16 4432 AGCcGACCCCACTGGGTATaTtCCatTgGGAaCAagCCTCCcAcAGCTACAGatacAcTtGctCCTg  
 31 4369 TGGgGGTGGCACTGGaTATgTtCCcTtTaGtACAGcTCTtCTACAGTaTcGgaggCAagTATaCCTa  
 18 4426 TGGcagTgGtAcAGGGgTgtaCagggtacAttCcatTgggTggggcgtTcaAataCagTggTggaTg  
 25 con tGGcgGtcGtaCtGGgtaTggttcC-tgggaAct-ctccc-ctacagctactaatcacg-gcc-cctg  
 BE11-GAAGCXKXCCCAAGCCXGKXAX-BE11  
 BE12-TATATACCTTGGGAAGCXKXCCCAAGCCXGKXAX-BE12  
 021-TGGGGGTCTACTGGCTATGTTCCCTTACAAACTCTCGAAaACCTCTTATTACTAGTGGGcCTATGG-021  
 30 6 4685 CtCGTCCtCCGtGgTgG7GGAGCCTG7gGCCCCCTTgGATCCaTcTATTG7GCTtTAA7TGAaGaa  
 11 4676 CACGTCCgCAcGTGcTGTGGAGCCTGTGGCCCTTCCcGATCCcTcCaATTG7GCTtTAA7TGAaGaa  
 33 4472 TACGTCCtCCgGTtACTGTAGAcCCTGTGGaCCTTtaGActCgTCTATAGTGTcATTAA7aGAAGAA  
 16 4500 TAAAGACCCcCtTtTaCaGTAGAtCCTGTGGcCCTTtGAtCCcTCTATAGTtCTTAA7TGAaGaa  
 35 31 4437 TTAGACCAcCAgTTagCaTtGACCCCTGAGGcCCTTGGACCCCTCTATAGTAAgTcTGTGTGAAGAA  
 18 4494 TTgGCCtAcAGTtccCcaGtggTtTaTgaACtGtGGgCCCCaCagaccAtcTaTGTtAcAttA  
 con t-cGtCCtCagttac-gtaGagccTgTtGgcCctt-gGa-cCtCtatagtgtcttTa-TaagaGaa  
 40 BE26-CGKCCXCCGGXKXACGXAGAKA-BE26 J324-TCATATGTGTCN7TAATNGAAGAA  
 BE27-GKCCXCCGGXKXACGXAGAKA-BE27 G22-GGATCCCTCTATTGTGTCCTTAA7TGAAGAA  
 BE28-XCCXCCGGXKXACGXAGACGXKXGACCCXKXAG-BE28  
 021-CTCGTCCCTCTGTGGTGGTGGAGCCTGTGGCCCTTCCGGATCC-021

6 4753 TCgGcAAtcATTAAcGcAGGgGcGcC TGAAaTtGtGCCcC TgCAGcAGGTGGgTTTAC  
 11 4744 TctGCTATTATTAAcGcGCTGcAC TCAGgtgGtaCCcC TACAGcAGGTGGgTTTAC  
 5 33 4540 ACaAGTTTTATaGAGcAGGTGCACCA GccCCaTctATTCC TACACcATCAGGtTTTga  
 16 4568 ACTAGTTTTATTGATGCTGGTGACCAcAaCatCTGtaCctTcCATTCcCagatgtATCAGGaTTTTag  
 31 4505 tCTGGAaTTgTTATGTTGGTGC cctGGCTCCTaTaCcacacCTCCaCaCATCTGGGTTTTga  
 10 18 4562 ataGaggactccagTGTgGtTaCatccaggTGCaCTAGgCtAcgTtTACtggcAGcTCTGGGTTTTga  
 con -ctgggtattatt-atGctGgtGcCacca-ctgctgc-at--c--ccctcct-caccatctGGgTTT-a  
 TCTAGTNTTATTAAATGCAGGTGCACc-JJ22  
 BE5-GCAAXCAXXAACGcAGGGGGCC-----XGAA-BE5  
 15 BE6-GCAAXCAXXAACGcAGGGGGCC-----XGAAAXXGAGCC-BE7  
 05-GTACCCCC TACACAGGTGGCTTTAC  
 022-TCGGCAATcATTAAcGcAGGGGGCC TGAAaTTGTGCCcC TGCACAGGTGGGTTTAC-022  
 6 4812 aATtACATccTCTGAaAcCAcTACcCCTGCAaTaTtGATGT ATCaGTT ACtAgTCACACTA  
 20 11 4803 TATAaCATCATCTGAaTcGACTACACCTGCTATTtTaGATGT gTCTGTT ACcAaTCACACTA  
 33 4599 TgTTACTACATCTGCAGATACTACACCTGCaATTATTAAATGTTtTcATCTGTTggggAgTcATCTATTc  
 16 4636 TATTACTACTTcaAcTcGATACcACACCTGC tATATTA GATattAaTAACTGTGTTA  
 25 31 4570 cATTgCTACaAcTcGcAGAcACaACACCTGC aATTTTA GAT gtaCAaAgTGTGTTA  
 18 4630 tATaaCatCtGcGgtacaaACTACACCTGCGgtTTTggatatacacaccttcgtataacctCtgggtcTA  
 con taTtAc--CatCtgcag--ACTACaCCTGCaatttTt-atgt--catctgtt--tac-aat--ta-Ta  
 05-TATAACATCATCTGAATCGACTACcCTGCTATTtTAGAGGT GTCTGTT ACCAaTCACACTA-05  
 30 022-AATTACATCCCTGAaAACCACTACCCCTGCAATTTGGATGT ATCAGTT ACTAGTCACACTA-022  
 6 4874 CtACTA GTaTaTTTgAAATCctGtCTTACAGAAcCTcTGTGAaCAaACCCCAACCCACCcGTG  
 11 4865 CcACTA GTgTgTTTTcaAAATCCcTgTTTACAGAAcCgTCTGTAAATAGcCCCAACCAcCTGTG  
 35 33 4667 aaACTATTtCTACACATttaaATCCCAcATTTACTGAACATCTGTAAATAGcCCCAACCAcCTGTG  
 16 4692 cTACTgTTaCTACACATaTaAATCCCACTTTcACTGACcCATCTGTATTGcAGCTCCCAACCAcCTGTG  
 31 4623 gCACACATgAaAATCCTACTTTTACTGATCCATCTGTATTGcAGCTCCcACCAcCTGTG  
 40 18 4698 tttccacaacCAatttTaccAACTCTGcATTTcTGATCCgTCCaTaTgAgTtCCCaAaACTGg  
 con ctactatta-taca--TaaaAATCC-ac-TTtaCtGAaCAtCtGtTaatacAGcctCcacacCtGc-  
 05-CACTA GTGTGTTTTAAAAATCCcTGTTTACAGAAcCgTCTGTAAATAGcCCCAACCAcCTGTG-05  
 022-CTACTA GTATATTAGAAATCCCTGCTTTTACAGAAcCTTCTGTAAACAAACCCCAACCAcCTGTG-022  
 45 027-CTGCA

[illegible]

6 5143 TCCACtCctCaaCGGTTaattAcaTAT GATAACCCCTGTATATGAA GGgGagGATG  
 11 5134 TCCACgCcaCagCGattggTAATCTAT GACAAACCCCTGTATGAA GGAAGaATG-05  
 5  
 33 4932 TTTTtTAACatCgCCTcTAAACCTTATACATATGATAATCCTGCATcTGAAGGctTgAGccctGaaG  
 16 4957 TTTTgTAACCaCTCCcAGTAAACCTTATTACATATGATAATCCTGCATATGAAGGTATAGATCtgGaaA  
 18 4879 TTTCTTTAgTgCTCCAAaagAcCTAAATTACATATGATAAACCCTCCCTATGAACtTgTAATGCTGaaG  
 10  
 18 4963 GTTCTTAcagTCCAtcctcttTAATTACATATGacAAACC gCctcttG  
 con tttttt-accactcctta-taacttATtacatatGATAAcCCtgcataatgaagt-taga-gc-gatg  
 05-TCCACGCTCACGCGATTGGTAACCTAT GACAAACCCCTGTATGAA GGAAGaATG-05  
 022-TCCACCTCCcCAAGCCTTAATTACATAT GATAACCCCTGTATGAA GGGGAGGATG-022  
 027-TTTTTTAACATCCGCTCATAAACCTTAAACATATGATAAACCCTGCATTTGAAGCCTTGACCCCTGAAG-027  
 15  
 6 5198 TtAGTgTACAATTTAgTCATGAtCTTA TaCACAATGCACCTGATGAGGctTTTATGGaATa  
 11 5189 TaAGTTTACAATTTACcCATGAgTCTTA TcCACAATGCACCTGATGAAGcATTATAGCATtT  
 20  
 33 5000 AcACATTACAATTTcaACATAgTGATA TAtcaccTGCTCCTGATCCTGACTTtTcAGATATT  
 16 5025 AtACATTATActTTTCTagTAsTGATAatagtaTTAATATAGCTCCaGATCCTGACTTtTtGATATA  
 31 4947 AAcCTTTTAACTCTTTC caATAcATcGTCAAAATATAGCCCTGATCCcGACTTtTAGATATT  
 18 5013 AgcCTgTggACacTaCattaacattTgatcCtCgTAgTgatGtCCCTGATTcCaGATTtTtTGATATT  
 25  
 con a-acttTAcAattTAc-cataaattAast-ctctttaaataatGctCCGATcc-CacTTtTgGATAtT  
 05-TAGCTTTACAATTTACCCATCAGTCTA TCCACAATGCACCTGATGAAGcATTATAGCATATT-05  
 022-TTAGTGTACAATTTAGTCATATCTTA TACACAATGCACCTGATGAAGCCTTTATAGGACATA-022  
 027-ACACATTACAATTTCAACATAGTGATA TATCACCCTGCCTCGATCCTGACTTtTAGATATT-027  
 30  
 6 5260 ATtCgTtTgCAGAGACctGcAtTgCCTCCGAGCTGGcCTTGTCGGgTAcAGTCGcATTGgacAAAG  
 11 5251 ATTtGACTACATAGACCAGCTATsAGCTCCaGACGgGCTCTTGCGCTTTTACTGCATTTGGcAAAG  
 33 5062 ATTGCATTACATAGGCCtGCTATtAcATCTcGtaGacaTactGTGCGCTTTTACTAGAgTAgGCTCAAAA  
 35  
 16 5093 gTTGCTTTACATAGGCCAGcATtAACCTCTAGcGcTACTgGcaTTAGgTAcAGTAGAaattGCTAATAA  
 31 5009 ATAGcATTACATAGGCCCTGCTTACCTCacGtaGgAacACTGTtTAGATATAGTAGACTAGTAATAA  
 18 5081 ATccgctACATAGGCCCTGCTtTAACATCCaGcGtgggAGCTGTTCcGctTtTAGTAGATtAGGTCaag  
 con aTg--tTAcATAGgCCTcGtaT-ac-TCC-G-cGtggtactgT-G-T-tAGTAgaat-GGtcAa--  
 40 JJ24-TACATAGGCCCTGCTATAACNTCCAGNCGTGGNNtTgCGNtTtAGTAGA-JJ24  
 05-ATTAGACTACATAGACCAGCTATAACGCTCCAGACGGGCTCTGTGCGCTTTTAGTCGATtGGGCAAG-05  
 016-ACGCTGCTTTTAGAGACAGTGTCAAAA-016  
 022-ATTGCTTTTGACAGACAGCTGCCATTGGCTCCcGAGCTGCGCTTGGCCGCTACACTGCCATTGCAACAG-022  
 027-ATTGCAATTAGGCCCTGCTATTACATCTCGTAGACATACTGTGCGCTTTTAGTAGAGTAGGTCAAAA-027  
 45 028-GTAGACATACTGTGCGCTTTTAGTAGAGTAGGTCAAAA  
 50  
 55

```

5 6 5328 GGGGTCATGCAACACCGCAAGCGAAAGCAATAGGGCCCGGATCATTATT
11 5319 GGGGTCATGCAACACCGCAAGCGAACAATAGGCGCCGCGATCATATT
5 33 5130 AGcCACACtAaAACTCGCAGTGGtAAACAAATGGAGCTAGATACATATTATcAgGATTAAAGTc
16 5161 ACaAAACtACgACCTcAGGATcAGGAaAACTATcGGCGAAATAGGCAAGATATATGATGATcAGTc
33 5077 ACaAAACtTgGCGACtCGTAGTGGTgcTcATtTGtGGTcAAGGgTgCATtATtATGATtAGTt
10 18 5149 ggcAACtATGttLAcCGCAgCGGtAcceAAtAGTGGtCAGGGTcAcTtTATcATGATtAAgTc

con
-g-aacTATgcacATtCGcAGtGG-aaacATtAGGtGCTagc-TACATATTtAcctgataagta
05-GGGGCTCMTGACACAGCGAGTGCACATATAGCGCCGATCATATT (-05)
016-AGCCACACtATTAAGCTCGGAGTGGtAAACAAATGGAGCTAGATACATATTATcAGGATTAAAGTc-016
12-0GGGCTCMTGACACtCGGAGTGCAGCGAAAGCAATGGAGCGCCGATCATATT (-022)
027-AGCCACACtTAAACCTCGGAGTGGtAAACAAATGGAGCTAGATACATATTATcAGGATTAAAGTc-027
028-AGCCACACtTAAAGCTCGGAGTGGtAAACAAATGGAGCTAGATACATATTATcAGGATTAAAGTc-028

20 6 5381 TTtAGATATTTCCACcATtTgCAGAGCTGCGAGGAATAGAAATGCACCCCTCTtGTGG
11 5372 TtAgGAcATTTCAACAGtTtACAGAGCTGCGAGAGAAATAGAAATGCACCCCTCTAGTGG
33 5198 CTATTG TgcCtttAGACcAcacccTgCAAATGaaCAATtTgAATcAcAgcCTttCaTgtaact
16 5229 CTATTGATCTCGCAGAGAAAtagactTtAACTataAcAcCTtCtAcACAcTACACCTTCacat
25 31 5145 gTATTATCTCGCAGtGAAAGTATTGAAATGCAcCTTtAGgggCgTtGCACtACACCTTCcact
18 5217 CTATTgcaCCTtCcccaAGAtATtTGAGAcCTGAGCGTTTAG tATTCGC caAGggag

con
ctattgato-t-c--agacacat-taca-aggc-caa-g-aatc--aa-ctacccc--tcg---
(05-)TTCAGGACATTTCCACAGTTTACACAGCTGCGAG-05
016-CTATTG TGcCTTtAGACcACACCGTGGCAATAGCAATATGAAGTACACGCTTTACATGATCT-016
022- TTATGATATTTCACcATTtTGCAGCAAGCAG-022
027-CTATTG TGcCTTtAGACcACACCGTGGCAATAGCAATATGAAGTACACGCTTTACATGATCT-027
028-CTATTG TGcCTTtAGACcACACCGTGGCAATAGCAATATGAAGTACACGCTTTACATGATCT-028

35 6 5441 CTGCGAcaggATGAtACATtTGATATTATGCTGAActtTTtAGAcCTggCAtTAcCCCTAcCCAAcAC
11 5432 CTGCGAGAAATGACAGCTGTTTATTATGCTGAAcATTtTGAcCCTtAcceCTgACCCGTtCAACAT
33 5263 tCtAcTcTgCTtTAcTATGATtTATtAgTGG tTtTgATGATgTtTATGc TgaGtAtGT
16 5297 tCAAGCTcAcCTAcCTtTtATTATAGGA TtTAAcATtTtATGAGATgacttTATACAGA
40 31 5213 tGAATtGatggCTTtATGAcATtTTATGCAGa CAcTGATtTtCTGTGGATacACCTGcACACa
18 5273 gAcAATGA CTtTtTGATtATtATGCAGATgAcAtGAcceCTgCAcGTGcTgTAcCATCGcgtc

45 con
-t-a-a-a-a---tat-T-tAT-tATgT-acag-ac-acgatattc-tgc---taccctt-ccaa---
016-TCTACGCTGCTTATAGATATATAGGg TGTGAGATGATGTATGC TGACGATCT-016
027-TCTACGCTGCTTATAGATATATAGGg TGTGAGATGATGTATGC TGACGATCT-027
028-TCTACGCTGCTTATAGATATATAGGg TGTGATGATGTATTAGC TGACGATCT-028

```

6 5509 cCTGTTACAaatatatacagAtaCaTATtTeAcTtCCACACCTTAATACagTtTaCACAACCGTGGGGTAA  
 11 5500 tCTGTTACA cAgTcTtATCTtACGtCCACACCTTAATACCTTtCACAAtCGGGGTAA  
 33 5319 ggaTaaTgtAcaCaCccCAATgCAacCaCTCAcAgtaCgTtTgCAaCaacacgTACCaGCAATGTgt  
 16 5362 TAcTtTaCaACcCcGggtAcCaTctgttaoCCTCtACatCTTtTaCAGgtTATaTTCCTGTCAAAATACAA  
 31 5278 TAtTgtTTCCcCTtCtAcTgCTtCAgACgTCAcCAtCTgCTgTgTtTtGCTtATgTtACCTCAAAATACCA  
 18 5338 TAoTcccTCCtTgCattTtTtAaAtAtTtGgcCcaCTataTctTCTGCCCTtTtccTaTAgTAAATgTAA  
 con ta-tttt-ca---tctc--attca---tcacctac--c-ttatcagcct-tc-ca-tagtaatgtaa  
 016--GGATAATGTACACACCCCAATGCAACACTCATACAGTACGTTTGCAACAACACGTACCAGCAATGTGT-016  
 027--GGATAATGTACACACCCCAATGCAACACTCATACAGTACGTTTGCAACAACACGTACCAGCAATGTGT-027  
 028--GGATAATGTACACACCCCAATGCAACACTCATACAGTACGTTTGCAACAACACGTACCAGCAATGTGT-028  
 6 5577 CACCACAGTtCCATTGTCAcTtCCCTAA TGACctGTTTtTeCAaTCTGGcCTGATATAAATCTTCCCTA  
 11 5559 tACCACAGTtCCATTGTCAATcCCTAgTGACTGGTTTGTgCAGTCTGGgCCTGACATAACTTTTCCCTA  
 33 5387 CtATACCTTtAaAtacaGAtttgATAcTCTGTtTtAgTtCaGgTCTGATATACCTtTccCCTTtAtTt  
 16 5430 CAATtCCITTTgTGgttgCaATacAAATTCCTTtAgTtCaGgTCTGATATACCTtTccCCTTtAtTt  
 31 5346 CtGTgCCacTAAGTaCaGgtTtTtCAcATTCCCaTATtTtCtGGGCTGTgTACCTATagAgCATgCA  
 18 5406 CgGTtCCtTtAAcctCctcTtYggGAtgTgCctgTATaCaGgGtCCTGAT AttcatttAccATtCa  
 con ca-taCcttttaaat-tg-aTtgcataT-cCtgt-tt-tc-ggtCctGat-taccataacattt-cta  
 016-CTATACCTTTAAATACAGGATTtGATACTCTGTTATGCTGGCCCTGATATACCTTCCCTTTTATTT-016  
 027-CTATACCTTTAAATACAGGATTtGATACTCTGTTATGCTGGCCCTGATATACCTTCCCTTTTATTT-027  
 028-CTATACCTTTAAATACAGGATTtGATACTCTGTTATGCTGGCCCTGATATACCTTCCCTTTTATTT-028  
 6 5645 CTGCAcTATGGGAACACCCCTTtAGTCTGTGAACCTCTGCTTTACCTACAGGCCCTGTtTtTtATTACA  
 11 5627 CTGCAcTATGGGAACACCCCTTtAGTCTGTGAACCTCTGCTTTACCTACAGGCCCTGTtTtTtATTACA  
 33 5455 CcCacAtCTAGccCATtTgt TCCtATtCgCCTtTtTTTCCTtTtGACACcATtTgtTGTgGAc  
 16 5498 gaCcaAgCTccITCAATAA TCCtATtCtTcCagggTCTCACAAtaTACAAtaTtTGTGAT  
 31 5414 CctACACAGgtTtTtCCCAAT TCCtTtTGGCCCTtAcAaCgCCACAGgtGCTATTtTGTGTGAT  
 18 5472 CtacCctGtaTgGCCATtTgtatcaCcaCgGCCCTTgCctCtAcACA GTaTATgTgTaTAcAT  
 con ct-c-act-tgtg-ac-a-tttttagtCCAtagctCtGctt-tcC-caag-c-ctaTtTt-ttgtat  
 BE22-CCCAXGXKXACCCACGGCC- BE22  
 BE23-CCCAXGXKXACCCACGGCCXGCCXKXACACA-BE23  
 016-CCCACATCTAGCCCATTTGT TCCtATTTCGCCTTTtTtTtCCTTTTgACACcATTGTGTAGAC-016  
 027-CCCACATCTAGCCCATTTGT TCCtATTTCGCCTTTtTtTtCCTTTTgACACcATTGTGTAGAC-027  
 028-CCCACATCTAGCCCATTTGT TCCtATTTCGCCTTTtTtTtCCTTTTgACACcATTGTGTAGAC-028

6 5713 GGTCTCTGgaTTtTAATTGCACTCTgCATGGTatTTTGACGtAaAGCCCGTAAACGTATTCCCTTATT  
 11 5695 GGTCTCTGACTTcTATTTCATCTCTACATGGTAcTTTGACGcAGACGCCGTAAACGTATTCCCTTATT  
 5 33 5518 GGTCTCTGACTTTgCTTTACATCTCTAGTATTtTtATTTACGtGcaagCGTAAACGTATTCCATATT  
 16 5561 GcaGGTGACTTTtTATTACATCTCTAGTATTtTACATGTTTACGaaacGACGCTAAACGTATTCCATATT  
 31 5477 GggGGTGATTTTATTTCACcCTCTAGTATTtTATATGTTAaaAcgtCGACGTAAACGGTATcCATATT  
 10 18 5537 GGTacacATTaTATTTCgtgGCattaTATTATtTtaTtccTaagaaACGTAAACGTGtTccCATATT  
 con GgtgctgacTtttATTgcatCCtag-TatTat-Ttttaccgta-acgaCGTAAACGT-TtcC-TatTT  
 JJ25-CGTAAACGNTTCCCTATT  
 PCR2-CGTTTCCATATT  
 027-GGTGCTGACTTTGTTTACATCTCTAGTATTtTATTTTACGTGCGAGCGTAAACGTATTCCATATT-027  
 15 6 5781 TTTTcCAGATGT GCGCGCCTAGCGACAGCAGCATATATGTGCTCTCTCCtAACCTGTATCC  
 11 5763 TTTTACAGATGT GCGCGCCTAGCGACAGCAGCATATATGTGCTCTCTCCcAACCTGTATCC  
 20 33 5586 TTTTACAGATGTCCgTgTGCGCGCTAGTGAGGCCACAGTgTAcCTGCTCTCT GTACCTGTATCT  
 16 5629 TTTTcCAGATGTCTCTtTGCGCGCTAGTGAGGCCACGTCTACTTGCCTCTCT GTCCAGTATCT  
 31 5545 TTTTACAGATGTCTCTGTGCGCGCTAGTAcGAGGCTACTGTCTACTTACCACCT GTCCAGTGTCT  
 18 5605 TTTTgCAGATGgctTtTGTCGCGGCTAGtGacaaTAcGtATAtcTtCCACCT ccttCTGTGCa  
 25 con TTTTACAGATGtctctgtTGCGCGCTAG-GA--cCacAGTaTA--TgCCTCTCTc-gtccCTGTatCT  
 TTTTNCAGATGCTCTGTGCGCGCTAGTa-JJ25  
 PCR1-CAGATGCTCTGTGCGCGCTAGTG-PCR1  
 TTTTGCAGATG-PCR2  
 027-TTTTACAGATGTCCGTGTGCGCGCTAGTGAGGCCACAGTGTACCTGCTCTCT GTACCTGTATCT-027  
 30 6 5843 AaaGTTGCTGCCACGGATGCTtATCTTAcTcCGCACAAcATATTTATCATGCCAGCAGTCTAGACT  
 11 5825 AAgGTTGTGCGCACGGATGGgTATGTtAaCGCACCAcATATTTATCATGCCAGCAGTCTAGACT  
 35 33 5651 AaaGTTGTcAGCACTGATGAATATGTgtCtCGCACAAgCATtTATTATATGCTGGtAGTTCACACT  
 16 5694 AAgGTTGTAAAGCAGGATGAATATGTtGcAGGCACAAcATATATTATCATGCCAGGAAcTCCAGACT  
 31 5610 AaaGTTGTAAAGCAGGATGAATATGTtGcAGGCACAAcATATATTATCATGCCAGCAGTGTAGGCT  
 40 18 5670 AAgGTTGTAAAcAcCGATGAtTATGTgAcTcCacAaAgCATATtTTATCATGCTGGACGctCTAGaAT  
 con AaaGTTGTaaGcACgATGaATATGTtAc-CgcAC-AacATAT-TTATcAtGC-qGcAgTtCTAGACT  
 JJ26-GTTGTNANCACGGATGANTATGTtACTTCGCACAA-JJ26  
 PCR3-AAGTTGTAAAGCAGGATGAATATGT-PCR3  
 LCR1A-AAGTTGTAAAGCAGGATGAATATGT-LCR1A  
 LCR1B-TGCAGCCACAAcATATATTATCA-LCRB  
 LCR1B'-ACGTGGGTGTTGTATATATATAGTA-LCRB'  
 027-AAAGTTGTGACGACTGATGAATATGTCTCGCACAAcGATTTATTATTATGCTGGTATTCCAGACT-027  
 45 6 5911 tCTTGcAGtGGGACATCCTtATTtTtCCAtaAAA cggGcTAA C AAAA CtGTGTGTC  
 50  
 55

11 5893 CCTTCTGTGGGACATCCATATTactCTATcAAAAA agtTAA C AAAA CAGTCTAC  
 5 33 5719 tCTTCTGTGTGGCCTCCATATTtTCTATTAAAAAaCCTAGTAA CgtAAAAAAATATTGgTAC  
 16 5762 actTTCAGTGTGGCCTCCcTATTtTCTATTAAAAAACCTAACAAAT AAcAAAA TATTaGTc  
 31 5678 gCTTACAGTAgGCCATCCATATTaTtCcATTcctCAAAcCTGACAACTcTAAAAA TACTGTGTc  
 10 18 5738 attTAACtGTtGtAAATCCATATT TtaggggtcCTGcaggTgTggcAAATAgcgGataTc  
 con -cTtgC-GTtGgacATCCaTATTtctcAtsaaaaaactgctaata--caacaaaaa--tagtgtTc  
 JJ27-GTtGGACATCCATATTtT-JJ27  
 027-TCtTCTGTGTGGCCTCCATATTtTCTATTAAAAATCCTACTAA CGCTAAAAAAATATTGgTAC-027  
 15 6 5967 CAAAGGTGTcAGGATATCAATacAggGTaTTTAAGGTgGTGTTaCCAGATCCTAACAAaTTTGcATTg  
 11 5949 CAAAGGTGTcLGGATATCAATATAGaGTgTTTAAGGTaGTGTTgCCAGATCCTAACAAgTTTGcATTa  
 20 33 5784 CcAAAGTATCAGGcTgCAATATAGGGTtTTTAGGGTcGgTTTACCAGATCCTAATAaTTTGcATTt  
 16 5824 CtAAAGTATCAGGATTACAAATacAGGGTATTtAGaaTaCaTTTACCTGAcCCcAAATAgTTGGcTTT  
 31 5743 CcAAAGGTgCAGGATTACAAATATAGGGTATTtAGGGTtCgTTTACCAGATcCAAAcAAATTTGGaTTT  
 25 18 5800 CtAGGTtCtLGGATacCAATATAGaGTATTtAGGGTgCagTTTACCTGAcCCcAAATAAATTTGGcTTa  
 con CcAAAGGTgTCaGgAT--CAATATAGgGTATTtAgggT-c--TTaCCAGATCCTAA-AaATTGyaTtt  
 JJ28-CAATATAGGGTATTtAGGGTNCNGTTACC-28 30-AAATAAATTTGGATTN  
 PCR4-GTATATCCCAATAATCCCATGTTAA-PCR4PCR5-TTATTTAACCAAA  
 027-CCAAAGTATCAGGCTTCaATATAGGGTtTTTAGGGTCCGTTTACCAGATCCTAATAAATTTGGaTTT-027  
 30 6 6035 CCTGacTCgtCtcttTcGatCCCCAaCACAaCGTTTAGTATGGGCaTGcACAGGcTaGAGGTgGG  
 11 6017 CCTGAtTcaTCCCTgTTTGACCCcACTACACAGcGTTTAGTATGGGcGTGcACAGGgTgGAGGTAGG  
 35 33 5852 CCTGACACCTCCTTTTATAACCCcGATACACAGcATTAGTATGGGCaTGTGTAGGcTTGAAaTAGG  
 16 5892 CCTGACACCTCaTTTATAATCCaGATACACAGcCGcTgGTTTGGGcCTGTGTAGGTgTTGAGGTAGG  
 31 5811 CCTGATACaCTTTTATAATCCGAAAcTCAACGcTTAGTTTGGGcCTGTGTGTGTTTaGAGGTAGG  
 40 18 5868 CCTGATAcTagTATTATAATCCGAAACaCAAGcTTAGTgTGGGcCTGTGTGTGagTgGaaTTGG  
 con CCTGA-sC-tcttTcTataATCCtga-ACaCaCGtTAgTaTGGGcGTGg-sGG--T-GAggTaGG  
 CCTGACACCTCNNTTTATAAT-JJ30  
 GGACTGTGGA-PCR5  
 027-CCTGACACCTCCTTTTATAACCCGATACACAGcATTAGTATGGCCTGTGTAGGcCTTGAATAGG-027  
 45 50 55



6 6103 CAGGGGCAgCCaTTAGGtGTgGOTCTaAGTGGaCATCCtTtCTAAAtAAATATGATGATGtGtAAAA  
 11 6085 CAGGGGtCAaCCtTTAGGCGTtGTGTtAGTGGaCATCCaTtGCTAAACAAATtGATGATGtGtAAAA  
 5  
 33 5920 TAGaGGGtCAGCCaTTAGGtGTgGOTCTaAGTGGaCATCCtTtCTAAAtAAATtGATGATGtGtAAAA  
 16 5960 TCGtGGtCAGCCaTTAGGtGTgGOTCTaAGTGGaCATCCtTtCTAAAtAAATtGATGATGtGtAAAA  
 31 5879 TCGtGGgCAGCCaTTAGGtGTtAGGtATAGTGGtCATCCaTtTtAAATAAAATtGATGACACTGAAA  
 10  
 18 5936 cCGtGGtCAGCCtTTAGGtGTtGGcCTAGTGGgCATCCaTtTtAAATAAAATtGATGACACTGAAA  
 con --G-GGtCAgCCaTTAGGtGTtGG-attAGTGG-CATCC-TTattAAATAAAATtGATGACactGAAA  
 027-TAGAGGGGAGCCaTTAGGCGTtGGCATAAGTGGtCATCCtTtTtAAACAAATtGATGACACTGAAA-027  
 15  
 6 6171 AT tcaGGGgTGGTGGTAaCCCTGGaCAGGATAaAGGGTTAAATGATAGGTATGATTATAAACAA  
 11 6153 ATaGtGgtGGTATGGTGGTAaCCCTGGtCAGGATAaTAGGGTTAAATGATAGGTATGATTATAAACAA  
 20  
 33 5988 cCGGTAAcCaGTATcCTGGaCAaCCgGGTGTgATATAGGGAATGtTATCCaTGGATATAAACAA  
 16 6028 ATcGTAGTgctTATGCAaCAaATgCAAGTGTgGATAATAGaGAATGTATATCAATGGATTATAAACAA  
 31 5947 AcTCTAAATagaTATGCGgTgTgTcCTGGcactGATAATAGGGAATGTATATCAATGGATTATAAACAA  
 25  
 38 6004 gTCCcATgcccCaCgtCTaATgTtctgagGAcgtTAGGGAcaTgTgTgTgTATATAAGCaG  
 con at-ctAaTggtgTgctgTgTaatcctggtgagGAtaatAGgGaa--TgTatctaTgGATTATAaCAa  
 027-CCGGTAACAAGTATCTCGGACAACCGGGTGGTGATAATAGGGAATGTTTATCCATGGATTATAAACAA-027  
 30  
 6 6236 ACaCAaTATGcATGGtTGGaTGTGcCCcCCtTtTgGGcGagCATTGGGGTAAaGGTAAaCAgTGTaC  
 11 6221 ACcCAGcTATGTATGGTGGcTGTGtCCaCCgTtAGGtGAACATtGGGGTAAgGGTAcACAaTGTtC  
 33 6056 ACACAGTtATGtTtTAcTtGGaTGAAGCtCCaACaAGGGGAaCATTGGGGTAAAGGTgTgTGTAC  
 35  
 16 6096 ACAGAAATtGTCTTtTAAATGGTtGCAAAACCACTATAGGGGAaCAcTGGGGAAAGGATccCcatGTAC  
 31 6015 ACACAAcTGTCTTtTAcTtGGTtGCAAAACCACTATTGGAAGaGCaTGGGGTAAAGGtAgTCTTGTAG  
 40  
 18 6072 ACACAgTtATGtAtTtTgGGcTgtgCCcCtTgCTATtGGgGAaCAcTGGGGTAAAGGcAcTgCTTGTAA  
 con ACaCA-tTATGt-Ta-TGG-TGt---CCacCtataGGGGAaCAcTGGGGTAAAGGtAcctcctGTGAc  
 027-ACACAGTtATGtTtTACTTGGATGTAAAGCTCCAACAGGGGAACATtGGGGTAAAGGTGTGCTGTAC-027  
 45  
 50  
 55

6 6304 CAATACacCTGTACaggtGGTGACTGCCGCCCTTaGAACCTTATTACCAGTGTATACAGGATGGCG  
 11 6289 AATACcctCTGTACaAaaTGGTGACTGCCGCCCGTTgGAACCTTATTACCAGTGTATACAGGATGGCG  
 33 6124 TaaTGcAGcMcTgCaaTGATTTCTCCACCTTaGAACTTATAAJAaTACTATTATTAGGATGGCG  
 16 6164 CAATgtTGcAGTAAaTCCAGGTGTGTCTCACCATTAGAGTTAAATAAACACAGTTATTAGGATGGCG  
 31 6083 TAAGaaTGCTTaTACcCTCGGTGATTGTCTCTCCATTAGAATTAAAAAAATcAGTTATaCAAGATGGCG  
 18 6140 atcgcgTcCTTtATcCaGgGcGATTGcCCcCTTAGAaCTaAAaAAAcACAGTTtTgAAGATGGCG  
 con -asta-tgCgtac--cctggtGAtTG-CC-CC-TTAgAcTtAtaAacac-gTtATacAGGATGGCG  
 027- TAAATGcAGcACCTGCCAATGATTGTCCACCTTTAGAaCTTATAAAATcATTATTAGGATGGCG-027  
 6 6372 AtATGGTtGAcACAGGCTTTGGTGCTATGAATTTGGcTGAtTTgCAgACCAATAAATCaGATGTCCT  
 11 6357 AcATGGTtGAcACAGGCTTTGGTGCTATGAATTTGGcAGAcTTcAAACCAATAAATCGATGTCCT  
 33 6189 ATATGGTgGAcACAGGATTGGTtgcATGGATTTTAAACATTgCAGGCTAAATAAAGTGAATGTCCT  
 16 6232 ATATGOTTcATACGGCTTTGGTGCTATGAGcTTTACTACATTACAGGCTAAcAAAAGTGAaGTTCCA  
 31 6151 ATATGGTtGAcACAGGCTTTGGAGCTATGGAATTGACTgCTTACAAAGAcCTAAAGTATaTGTCCT  
 18 6208 ATATGGTgATAcGGGATATGGTGCcATGGAcTTTAgTaCaTTgCAAGAcACTAAATGTgAgGTACCa  
 con AtATGGTtGAcAGGcTtTGGTgctATGGAAtTTact-caTT-CA-gc-AtAAAA-gtGATGTCCT  
 ATATGGTtGAcACAGGCTTTGGTGCTATGA-36 37-CAATNCANcMAATAAAGTGAATGTCCT  
 027-ATATGGTgGAcACAGGATTGGTGTGcATGGAATTTAAaACATTGcAGGCTAATAAAGcATGATGTCCT-027  
 6 6440 aTTGAcATaTGcGcACTaCaTGAATATCCaGATTATTaCAAAATGGCTGCAGACCCaTATGGTGA  
 11 6425 cTTGATATTtTGcGAcTgtcTGCAAAATCCcGATTATTgCAAAATGGCTGCAGACCCcTATGGTGA  
 33 6257 aTTGATATTtTGcGcAgTAcATGCAAAATCCAGATTATTaAAAAATGcTAgTAgCCTTATGGTGA  
 16 6300 cTGGATATTtTGAcATcTATTGTGCAAAATCCAGATTATaTTAAATGGTgCAGAcCAATATGGCGA  
 31 6219 TGGGAcATTtGTAAcTCTATTGTAAATATCCAGATTATcTTAAATGGTGTcGAGcCAATATGGCGA  
 18 6276 TTGGAAATTtGTcAgTCTATTGTAAATATCCcGATTATcTAcAAATGcTGCAAGAcCTcTATGGGGA  
 con -T-GAATtTGtGg--cTAttTG-AAATATCCaGATTATTa-AAATGgtcgcAG-CC-TATGGTGA  
 NTNGATATTTC-TJJ37 JJ38-AAATATCCGATTATTTHAAATGG-JJ38  
 027-ATTGATATTtTGcGAcTACATGCAAAATCCAGATTATTAAAAATGcATAGTAgCCTTATGGTGA-027

6 6508 TAGATTATTTTTTtctTaCGaAGGAACAAATGTTGCCaAGACATTTTTTAAcAGGGCTGGcgagG  
 11 6493 TAGATTGTTTTTtattTgCGAAAGGAACAAATGTTTCCaAGACACTTTTTAAATAGGGCGGTACTcG  
 5 13 6325 TAGTtTATTTTTtcttcttCGacGtGAACAAATGTTTGTaAGACACTTTTTTAAATAGGGCTGGTAcAt  
 16 6368 cAGcTTATTTTTTATTtTACGAAGGGGAACAAATGTTTGTaAGACATTtATTAAATAGGGCTGGTAcTcG  
 31 6287 TACATTATTTTTTATTtTACGAAGGAACAAATGTTTGTaAGGCATTTTTTAAATAGCAAGGcAGcG  
 10 18 6344 TtCcaTgTTTTTtTgTTCAGcGtGAGCAgcttTTGctAGGCATTTTTTgAATAGcAGGtAcTa  
 con tag-tTATTTTTtattTaCGaaggGAaCAaaTgTTTG-tAGaCAcTtTtTtAAAtAGggCtGGtactg  
 NO 86/05816-GAGG  
 027-TAGTtTATTTTTCTTTCTTCGACGTGAACAAATGTTTGAAGACACTTTTTTAAATAGGGCTGGTACAT-027  
 15 6 6576 TGGGGGAACCTGTGCCCTGATacaCTtaTaaTtAagGgtAGTggaAAAcGcaCgTCTGTAGggAGTACT  
 11 6561 TGGGGGAACCTGTGCCCTGATGACCTGTtggTaaAAGgggGTsatAAcAGatCaTCTGTAGtAGTAGT  
 13 6393 TaGgaGAggCTGtTCCcGATGACCTGTACATTAAAGgtTCaGgaACTACTGcCTtAttcaaAGcAGT  
 20 16 6436 TTGGTGAAaaTGTaCCaGcGAtTtATACATTAAAGGCTCtGGgTCTACTGcaAATTAGCcAGttca  
 31 6355 TTGGTGAAAcGgTcCTactGACTATATATTAAAGGCTCCGGTTCaAcGcACTTTAGCTtaAcCaGT  
 18 6412 TgGGTGAcAcTgTCTCcaateCTTATATATTAAAGGcAcAGGTatgctCtGCTcaccctGgcAgCtGT  
 25 con TgGGTGAA-ctGTgCCTgatgac-Tata-atTAAAGGctctggtactactgC-tct-tagc-Ag-agt  
 TGGGGGAACCTGTGCCCTGATACATTATAATTAAAGGCTAGTGGAAATCGCACCTCTGTa-W086/05816  
 LCR2A'-GGACAACCAZTTTCCCCCAATTATT-LCR2A'  
 LCR2B-CAGATCATCTGTAGCTAGTAGT  
 30 LCR2B'-GTCTAGTAGACATCGATCATCA  
 LCR3A-ATTTATACATTAAAGGCTCTGGGTc-LCR3A  
 LCR3A'-AAATATGTAAATTCGGAGACCCAG-LCR3A'  
 LCR3B-TACTGCAAAATTAGCCAGTTCa  
 LCR3B'-ATGACCTTTAAATCGGTCAAGT  
 LCR4A-CCTTATATATTAAAGGcAcAGGTAT-LCR4A  
 LCR4A'-GAATATATAATTTCGGTGTCGATA-LCR4A'  
 LCR4B-GCCTGCTTCACCTGGCAGCTGT  
 LCR4B'-CGGACGAAGTGGACCTGCGACA  
 35 027-TAGGAGAGGCTGTTCCCGATGACCTGTACATTAAAGTTTCAGGAACtACTGCTCTATTCCAAGCAGT-027  
 40  
 45  
 50  
 55

6 6644 ATATATGTTaAcAcCcCgAGcGGCtTtTGGTGTCtCtGAGGCaCAATTgTTTAAATAAGCCATATTG  
 11 6629 ATTATGTaCAtAcCCTAGTGGCTCATTTGGTGCTTCaCAGGCTCAATTATTAAATAAGCCATATTG  
 5 33 6461 gcTTTtTTTCcAcCtCCTAGTGGaTCAATGGTtACTTccGaatCTCagTTATTAAATAAGCCATATTG  
 16 6504 AaTTATTTTCCTACACCTAGTGGtTcATGGTTAGcTcGATGccCAAAATTGAAATAAGCCtTATTG  
 31 6423 AcAtACTTTTCCTACACCTAGcGGCTCcATGGTtACTcCaATGCaCAAAATTtTAAATAAGCCATATTG  
 10 18 6480 gtgTAtTcTCCGctCtCCaAGtGGCTcTAtGTTACcTcGAGctccCagTgtTTTAAATAAGCCATATTG  
 con attTattttccc-aCaCctAGtGGcTctAtgGTTaC-TcTGA-gC-CAatTatTAAATAAGCCATATTG  
 JJ39-GTTACNTCTGANGCNAATTATTAAATAAGCCATATTG  
 15 AT-LCR2B'  
 AA-LCR3B'  
 TTA-LCR3B'  
 GT-LCR4B'  
 CAC-LCR4B'  
 027-GCTTTTTTCCCACTCCTAGTGGATCAATGGTtACTTCCGAATCTCAGTTAITTAAATAAGCCATATTG-027  
 6 6712 GCTaCAAAAaCcCaGCGaCATAAaATGCTATTGcTGGGGaAAcCAactGTTGTTACTGTGGTAG  
 20 11 6697 GCTtCAAAAaGcTcCAGGGaCATAAaATGGTATTGcTGGGGaAAcCAactGTTGTTACTGTGGTAG  
 33 6529 GCTACAAcGtGCACaAGGtCATAATAATGGTATTGTTGGGGaAAcCAggtATTGTTACTGTGGTAG  
 16 6572 GCTACAAcCaGCaAGGGcCACAATAATGGcATTGTTGGGGaAAcCAggtATTGTTACTGTGGTAG  
 25 31 6491 GaTgCAACGtGcTcCAGGGaCACAATAATGGTATTGTTGGGGaAAcCAgTATTGTTACTGTGGTAG  
 18 6548 GtTaCAaaagGCaCAGGGtCAtAaCaATGGTgTTGcTGGCaTAAcCaATTATTGTTACTGTGGTAG  
 con GcTaCAaa---GCaCagGCaCAtAA-AATGGtATTGcTGGGGaAAcCAaTAAATTGTTACTGTGtGaG  
 GCTACAAANNNGCACA-JJ39 J41-AATGGTATTGTTGGGGTAATCAATTATTGTTACTGTGGTAG  
 C6-GCHCAGGGWCATAAATGG-C6  
 30 c1-CTGTGGTAG  
 c7-CTGTGGTTG  
 c8-CTGTGGTAG  
 c10-CAGTTGTAG  
 c11-CTGTGGTTG  
 c12-CTGTGGTTG  
 c13-CTGTGGTAG  
 c14-CTGTGGTAG  
 c15-CTGTGGTAG  
 35 027-GCTACAAcGTGCACaAGGTcATAAATAGGTATTGTTGGGGCAATCAGGTATTGTTACTGTGGTAG-027  
 40  
 45  
 50  
 55

6 6780 ATACCACACGGCAGTACCAATGACATATG TGCATCCTGaaCTAaATCTTcCAGTACACc  
 5 11 6765 ATACCACACGGCAGTACCAATATGACACATG TGCATCCTGtCTTAAATCTgCTACATACAc  
 33 6597 ATACCACtCGCAGTACtAAATATGACtTTATG ceCacAGTaaCTAgTAcAgTACATATAAA  
 16 6640 ATACCACACGGCAGTACCAATATGCTtTATGCGCT GcCATATCtCACTTcaGAAACtACATATAAA  
 10 31 6559 ATACCACACGtAGTACCAATATGCTtGtGCT GCAATtGcCaAcagTgATGATACATATAAA  
 18 6616 ATACCACtCCcAGTACCAATtTaaCaATGCTGCTtctcACagTctcctgtacTgggcCATATgAT  
 con ATACCACACcGcAGTACCAATaTgaCatTaTGtGct--tgCa--g--aacta--ag--tactacATataaa  
 ATACC -J341 C16-CATCCGGAACTACACtCTTCA-C16  
 ATACCACACCGAGTAC-C1 C17-TCGTGTCTAAATCTGTCTACA-C17  
 15 ATACTACACCGAGTAC-C7 C20-CACACAAGTAACTAGTGCAG-C20  
 ATACCACtCGCAGTAC-C8 C23-CAGTCTCCTGTACCTGGG-C23  
 ATACTACTCGCAGTAC-C10 C31-TTGCAAAcAGTGTACTACATT-C31  
 ATACTACCCGTAAGTAC-C11  
 ATACTACCAAGACAC-C12  
 ATACTACTAGAACAC-C13  
 20 ATACCACCGTAGTAC-C14  
 ACACTACCCGAGTAC-C15  
 027-ATACCACtCGCAGTACTAATATGACTTTATG CACACAAGTAACTAGTGCAGTACATATAAA-027  
 6 6842 AATTtCtGATTATAAGAGTACATGCGcCATGTGGAGAGTATGATTACAAtTTATTTTCAATTATG  
 25 11 6827 AATTcAGATTATAAGGAATACATGCGcCATGTGGAGAGTtTGATTACAGTTTATTtTTCAATTtTG  
 33 6559 AATgAAATTTTAAAGAAATAtATAAGACATGtGaaGAATATGATcTACAGTTTgTTTTTCAACTATG  
 16 6705 AATAcTAaCTTTAAAGAGTAcTAcGACATGGgGAGGAATATGATTACAGTTTATTtTTCAACTtTG  
 30 31 6624 AgTAgTAaTTTTAAAGAGTATTAAAGACATGCTGAGGAATtTGATTACAAtTTATATTTCAAGTATG  
 18 6684 gcTAcCAaTTTTAAAGAGTATAgcAGACATGtTGAGCAATATGATTtGcAgTTTTATtTTCAAGTtTG  
 con aaTactaAtItTAA-gAgTA-ata-GaCATGt-GAgGAATATGATtCaAgTTTTATtTTCAaAT-TG  
 027-AATGAAAAATTTTAAAGAAATATAAAGACATGTTGAAGAAATATGATCTACAGTTTGTTTCAACTATG-027  
 35 6 6910 TAGCATACATTgTCTGCTGAAGTAAATGGCCTATATtCACACAATGAATCCcTCTGTTTGGAGAGCT  
 11 6895 TAGCATACATTATCTGAGAGTcATGGCCATATAcACACAATGAATCCtCTGTTTGGAGAGCT  
 40 33 6727 CAAAGTTACCTTAACTGCAGAGTtATGACATATATcATgCTATGAATCAGaTATTtTGAAGAGT  
 16 6773 CAAAAAaACCTTAACTGCAGAGTtATGACATATAcATtCTATGAATtCaCATATTtTGGAGAGCT  
 31 6692 CAAATATACATTATcCTGCAGAGCaTATGACATATCTTCAcAGTAAATcTgCTATTtTGGAGAGT  
 45 18 6752 LaetAtAcTtTAACTGCAGATgItATGtCcTATATTCAAGTAAATAgcagTATTtTGAAGAGT  
 con -AaaaTtACATtTa-CTGCAGAGTtATGAGC-TAtAttCA-actATGAATccc-cTATTTtGAG-GA-T  
 027-CAAAGTTACCTTAACTGCAG-027  
 50 55

6 5978 GGAACCTTTGGGTTATCGCCTCCcCcaAATGGTACATaGaaGATACcTATAGgTATGTgCAGTCACAG  
 11 5963 GGAACCTTTGGGTTATCGCCTCCcCcaAATGGTACAcTgAGGATACtTATAGaTATGTaCAGTCACAG  
 33 6795 GGcAaTTTGGGTTTaaCACCTCCtCCAtctGtTAgttTaaAGGATACcTATAGGTTTGTtaccTCTcCAG  
 16 6841 GGAATTTTGGGTTTaaACCTCCcCaggAGGcACacTAGAaGATACtTATAGGTTTGTt aaCCcCAG  
 31 6760 GGAATTTTGGGTTgacCcaCacCtCCcCAGGTtCTTTGGAGGATACcTATAGGTTTGTtaccTcCAG  
 18 6820 GGAACCTTTGGGTTtCCcCccCccCaaCtactagTTTGGtGGATACaTATcGtTTGTtacaTCTggtt  
 con GGaAcTTTGGtTtTa-c-cCtCCcCcaactggtac-tT-gagGATACcTATAgTtTGTtaca-tCacag  
 C21-TTGT AACCCAG  
 C34-TTGT AACCCAG  
 C35-GTTTGT AACCCAG  
 6 7046 GCCATTACCTGTCAaAAgCCCAcTcCTGAAAAGGAAAAGCcaGA TCCCTATAAGaAcTtAGTT  
 11 7031 GCCATTACCTGTCAgAAACCCACACCTGAAAAGGAAAAGCagGA TCCCTATAAGgataTgAGTT  
 33 6863 GCtATTACgTGCAAAAAaCagTACCTCCAAAGGAAAAGGA CCCCCTTAggtAAATATaCAt  
 16 6906 GCAATTgCtTGCAAAAAaTaCACCTCCAgcaCCTAaaGAAGATgATCCCCCTTA AAAAAATaCAt  
 31 6828 GCcATTACaTGCAAAaAacTGcCCcCcaAaagCCcAaGGAAGAT CCaTTTA AAgATAtgtaT  
 18 6888 GCtATTACcTGCAAAAggaTGCTgCaCcggtgaaAataAgGAT CCcTaT gAtaAgTtaaeGt  
 con GCcATTAcCtGTCAaAAacCct-cacCtc-aagggaaAaggaaGAt---cCccttaaa-eA-t-ta-tT  
 GCAATTGCT-C21 C18-CATACACCTCCAGCACCTAA-C18 J345-T  
 C19-GGATGCTGCACCGGTGA-C19  
 C22-AAAAACAGTACCTCCAAAGGA-C22  
 C27-TTTTGTGTCAGAGGTTTCCT-C27  
 C24-CACACCTGAAAAGAAAACAG-C24  
 C28-GTGTGGACTTTTCTTTTGTTC-C28  
 C25-CTCCTGAAAAGGAAAAGCCA-C25  
 C26-GAGGACTTTTCTTTTCGGT-C26  
 C29-CCAAAAGCCCAAGGAAGAT C-C29  
 C30-CAAAAAGCCCAAGGAAGAT C-C30  
 C32-CAGAAACCCACACCTGAAAAGAA-C32  
 C33-AGAAACCCACACCTGAAAAGAA-C33  
 GCAATTGCT-C34 O23-GGA TCCCTATAAGGATATGAGTT  
 GCAATTGCT-C35 O15-GGAT CCGCTAT GATAAGTTAAAGT

6 7110 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTGATCAGTATCCCTGGACGCAAGTTT  
 11 7095 TTGGGAGGTTAACTTAAAGAAAAGTTTCAAGTGAATTAGATCAGTTTCCTCCCTGGACGCAAGTTT  
 33 6927 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTGATCAGTTTCCTGGACGCAAGTTT  
 16 6973 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT  
 31 6892 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT  
 18 6952 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT  
 con TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT  
 TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT  
 15 C2-GATCAGTTTCCTGGACGCAAGTTT  
 C3-GATCAGTTTCCTGGACGCAAGTTT  
 C7-GATCAGTTTCCTGGACGCAAGTTT  
 015-TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT-015  
 023-TTGGGAGGTTAACTTAAAGAAAAGTTTCTAGTGAATTAGATCAGTTTCCTGGACGCAAGTTT-023  
 20 6 7178 TT gTT aCAAACTGGATATAGGACGCGTCT  
 11 7163 TTA TT gCAAAAGTGGATATAGGACGCGTCT  
 33 6995 TTA TTAcagGcAggtctttaaagcAaaAcctaaacttaaACctgcAGcccccaCAtCcaCCcgCA  
 16 7041 TTA CTACAGCAGGATtgaAGGcCaACCAAAATTTacAtAGGcAaAGCgaagGCTaCAGCCcCaC  
 25 31 6960 TTA tTACAGGcAGGATatAGGcCagctCCtAAATTTAAAGCAGGtAAAGC TagTgCACCC t  
 18 7020 TT ggTtCAGGcTGGAtTgcGtcgcaagCCcAcCaTaggcCctcGcAAAGC T tctg  
 con TTa--ta--agcaggattgagggcacaaccacaaataaa-a-cagcgaag-gatatag-gcacc-cct  
 30 015-TT GGTTcAGGCTGGATTGCTGCTCAAGCCACCAATAGCCCTCCGCAAGC T TCTG-015  
 023-TTA TT GCAAAAGTGGATATAGGACGCGAGCT-023  
 35 6 7209 CTatTGTACAGGTgtTAAAGCCGCCCTGTGTtTcCAaagCCTTgTgCtGCCCCAAAGCTAAAGCGegCC  
 11 7194 CTgCTGTACAGGTaTAAAGCCGCCcAGCTGTgTctAAagCCTTcACAGCCCCAAAGCgAAACGTaCC  
 33 7059 CaTgTCTgCAaaacgcAaaaggtttaaasAATAACAcTtTgTgtaattgtgtTAtgtTGTgtTtgg  
 16 7108 CcTCATCTACCTcCTACAACTcGCTAAAGC cAAAAACgTAAGGctGTAA CtsTGTGTGTGA  
 40 31 7021 CagCATCTACCACTACRcCaGCAAAAGCTaAAAAAAC TAAAAAGTAAtgtGatGTGTATGTaTaCaT  
 18 7075 CtcCATCTgCCACTAC gctctt TAAA cCTGcCAAGCgt  
 con Ct-catcTAc-actscaaa---c---a--t-sat-as-gtaa-ctg-a-c-c-t-a-C-tgta--tcc-  
 45 015-CTCATCTGCCACAC GCTCTT TAAA CTTGCCAAGCT-015  
 023-CTGCTGTACAGGTATAAAGCCGCCcAGCTGTGTtTcCAaagCCTTcACAGCCCCAAAGCgAAACGTACc-023

```

5      6 7277 AAAACtAAAgGTAATATATGTGTGt      ataTgTaeTGTT
      11 7262 AAAACcAAAAaGTAATATATGTGtGtcagTGCTGTGTGT
      33 7127 TctTGtctATgTactTtgtgTGT      TGCTGTGTGTGtTgtTGT
      36 7167 TgtTGaaTtaAGTGT      TGTGTGT      ATATGT
10     31 7088 GGTGTGctgTgTGCTAAGGTGTGTGTgtatGT      ATATGTGTGTGTttgtgtgtTATATA tg
      18 7113 GTG      cGTGTAAGTGC      caGgaagtaAATATGTGTGTgt gtatataTATATAcat

con      -t-tctataagtgat-tgtttgtg-----tgtgttgtgt-tatgtgtgtgt-----tatata---
015-GTG      CGGTGACGTGC      CGGGAAGTATATGTGTGTGTGTGTGTGTATATATATACAT -015
15     023-AAAACCAAAAGTAAATATGTGTGTGTGTGTGTGTGTGT-(023)

      6 7313      AT
      11 7302      AttTaPATG      T      TGTTGTA      gTGTGT
20     33 7167      TTGT TttTgTGATG      TGttacaaTgtATGTATGTGTACGtTaeGTGGTTG
      16 7199      TTGTATGTgcttGTATG      TG      CTTGTAAATATTAAGTTGTATGT      GTGTTTG
      31 7150 gTATGTATgTtTaTGATG CgTGTGT      aCTTGTAATAAT GtaTaeGTATGT      TATGTGTG
25     18 7164 cTATgtTgTGTT GATGTCcGTGTgttgtgtTGT TGTAT G      aTtgcatgTATG G

con      -tattgtatgttttgcatag-c-tgtgt-tgt-cttgatataatatgttgttatgttt-gtgtgtttg
015-CtATTGTGTGTGTTT GTATGTCTCTGTGTGTGTGTGTGT TGTAT G      ATTGCAATTGTATG G-015
30     023-      ATTATATATG      T      TGTTGTA      GTGTGT(-023)

      6 7315      ATATATGT
      11 7325      aTATGT TctTGT aTTGIG      TATATGT
33 7221 T      ttTATGTgtcTTGTttGtGtGcATGTtTATGTacttgt
35     16 7248 TATGTATG      gtaTAATAAA      cacGGGTATATG
      31 7209 TATGTATGctatgtatGTGTAATAAAtgtgtatacctgtgtgtgtGTGTATGTCTcctTataTAc
      18 7221 TATGTATG      gtTGTt      gtTGTATGTATGTtTcATGT

40     con      tatgtatg-----tgttaataaa-----ttatgt-ttcttgtt-gtgtgtatgct-tatgtata-tat
015-TATGTATG      GTTGTG      GTTGTG      CTTCGTGTGTGTGTGTGTGTGTAT -015
      023-      ATATGT TGTGTGT ATTTG      TATATGT(-023)

```



6 7323 GT gATGTAACGT  
11 7351 GTAtATGTTTGATATATGT GTAT gttATGTA TGT  
33 7262 cagtttccGTGTTGATGTATATGTTtaataaacattgtTGATATTgtTaaActATTgtATGTA TGT  
16 7279 GTTtCAAAgTcTgtgGATCATATGT gtcATGCAacTAaaTAaacctatT  
12 7277 AccctaTtagtaaacatactATAcTAtTttataAACTATTGTctcTactTgtTcctAcTgttctCTgc  
10 18 7257 AattgtTggtatgtggcattaaAaAaaTatgttttgggtTctgTgtgTtaTgtgggtTgcgcCTag  
con a---tat-tgtttgtgatat-ataatatagaacatbtgctttttatgtgaattat---Tatgtactgt  
015-ATTGTGTGGATGTGGCATTAAATAAATATGTTTGTGGTCTGTGTGATACGIGGTTGGCCCTAG-015  
15 023- GTATATGTTTGATATATGT GTAT GTTATGTA TGT-023  
6 7336 TATGT aTATGT CTgTGTGTGTtctGTGTGtaatgtaAgTATTTTGTGTAATGTGTATGTgCTT  
11 7386 TATGTgtTATGTatGTCTGTGTtTaaGTGT gTATATATTGTGGAAAGTGATGTATGTGT  
20 33 7329 TATGT AtatggggtgtaccTatatGAGTAagGagTtGTATTGCTtGccctacCctGCATTgc  
16 7331 gttTCAacAcccCActaattTgtTgtTggtTgtTctATGTATAAaactataTtGctACAcTgcTgtTt  
31 7345 TcctCCcaaAtagtcATgtacTtattTctgcgcTataAaTTTAggTgTcaagcccaTgATAaaAGTtgtaCa  
25 18 7325 TgagtaacAactgcATTtTgtggtttTgtggatgggtTtTgcttTgtTgggctatataTgtctcTgtatTt  
con tatgttaa-aa-gt-attttgt-tttt-tgtgtgtaagtatbttaatttgc-taa-ttgtatgt-tttt  
015-TGAGTAACAACCTATTTGTGTTTGTGGTATGGGTGTGTCTGTGTGGCTATATATGTGCTGTATT-015  
30 023-TATGTTGTATGTATGTTTGTGCGTTTATGTGT GTATATATTGTGGAAAGTGATGTATGTGT-023  
6 7400 TaTGTGCAATAAACCAATTAcctctTgtTAcacCCGT gACtCAGTgctgttgccagcGTTTtTgT  
11 7450 TtGTGCAATAAACCAATTA TtatgtgtgtcCTGTATACACCAGTG actaaGTgtTGTt  
35 33 7390 aaTGTAcCTAccttAtTtTcccTATAtTgtAGtaCCTACATGTttaaGTattgCtttacCTTTGaca  
16 7397 tgttTtATAtaactaTAtTtTtAggcGCAGgcCCatTTTGTacCTtCAacGgaAttcggTtGaat  
31 7413 CccggTccgtTtTttgtCAATaaGctactTCATTtTgATTttatGcGagCCAtTTTAAaTcccTAACC  
18 7393 CaaGtTataaaactgcacAccttAcagatTCATTtTatccTAcacatccctCcaTTTtgcTgtgcAACC  
40 con tatgttcaa-aat-ataaccttata-t-tcc-tt-t-acat-cagtg-c-attttaa-Cggtt-act  
015-CAAGTATAACCTGCACGACCTTACAGTACCTCAATTTCTCAACCTCCCTCATTTTGTGTGCGCAAC-015  
45 023-TTTCTGCATAAACCAATTA TTATGTGTGTCTGTTTACACCAGTG ACTAAGTGTGTGT-023  
024-GAATTGCGTTCGAT

6 7466 TTGCAGCGCCctTaccscataagTaATATacaTgcAcaATATATATaTtTttgtTaaatTACTAT  
 11 7508 TTGCAGCGCCGgtTtTgtttggccTCAATAT TatatTATATATATttTgTaataTAcCTATACATATg  
 33 7458 TactAgTgtCCaTATtTgtacsaTtTCCtccattTgTATGcCTAaccgTtTtcggTtACTTgGCATac  
 16 7467 GcTtTtTgGcCaAaaTgTgtTtTttaaTgTCTATGcCaggaacTaTgTtTaAacCTTGTACGT  
 31 7481 GtTtTTCGGTTCGAttgTtTaaacaTgctAgTACaaCTATGctgatgcagtGtTtcGgGgTtTtTgGt  
 10 18 7461 GaTtTTCGGTTCG ctttggcTtaTgCctgTggTtTt  
 con -ttt-cgg-ccttat-t-ta-a-ttc-tataa-t-ctatgt-tatat-ttt-tt-T-actttgct-tt  
 015-GATTTCGGTTCG CTTTGGCTTATGTCCTGGGGTTTT-015  
 023-TTGCACGCGCCGTTTGTGTGCTTCATAT TATATTATATATATTGTAAATACCTATACATATG-023  
 024-GCTTTTGGCACAAAATGTGTTTTTTTAAATAGTCTATGTGAGCAACTATCGTTTTAACTTGTACGT-024  
 15  
 6 7533 aCttttatAtTTGCAACCGTTTTCGGTTCGCCCTTAgCATACACTTtCCaCaAAITGTtTAcAAC  
 11 7573 tTACCcCcccccAcTTGCAACCGTTTTCGGTTCGCCCTTA CATACACTTAcCTCaaAAITGTtTAcAAC  
 20 33 7526 aTACCcCtsTgsCAatTGGCAGaacAgTtaaTccTtTtCtttCCGCACTGctgtTgtTcTgTACTTgctg  
 16 7535 TCTCTG cTtgCcaTgGgtGccAaaTcccTgtTtTcCTgCCTGCACTG cTTgccaAcaTtcc  
 31 7549 TTCTCTG aaTAcTAgTTTTtGcCaacaTTCtggcTtgTagt  
 25 18 7496 cTgCacaatacagtagctgtgcaactattgcaaaactTaaTctTtTggGCaetgtTcCTacaTatTttg  
 con tt-c-ct-tt-catt-gcagcc--tttg-tt-ctcttate-T-cact--c-tcttct-tattata-c  
 015-CTGCACAAATACAGTACGCTGGCACTATTGCAAACTTTAATCTTTTGGGCACTGCTCCTACATATTTTGG-015  
 023-TTACCCCCCCCCCACTTGCACCGTTTTCGGTTCGCCCTTA CATACACTTACCTCAAATTTGTATAC-023  
 024-TTCCG CTGCGCATGCTGCCAAATCCCTGTTTTCTCTGACCTGCAGC CTTGCGCAAGATTC-024  
 30  
 6 7597 GTGTTTccTctTAACTCtATAtttTGTG CcAGGTACAcATTGCGCTGCCAAGTgtCTTGCCAA  
 11 7640 GTGTTTtgcTACTAATCCcATAT gTGTGTgCcaAGGTACAcATTGCCCTGCAAGTatCTTGGCAA  
 33 7594 caTtggcaTACataCCcTATgacattTgCagaaCagTtAAcctTtTCTTtCTgcacTgtTtTgtc  
 35 16 7598 aTtTgttTtTtACAcTgCacTatgtgcaACtActgAaTCAcTaTgTgCATTgtTgCataTAAaaTaaT  
 31 7589 tTCTGgctTACACaCCTTgccaCATATAAaccAgTCCaacTtTGCAATTATaCtAtgAatCatgtT  
 18 7564 aaCaattggcgCgCctCTTtggcgCATATAA ggGcaccTGgtATTA gtcATtTtCtGtcc  
 40 con -t-ttt-ta-ca-tcTtatat--tt-taa-ccaa-g-aca--Ttgc-tt-c--aatt-ttt--a-  
 015-AACAAATGCGCGCGCTTTTGGCGCAATAA GCGGCACCTGCTGATA GTCAATTTGCTGTGTC-015  
 023-GTGTtTTCGACTAAATCCCATAT G-023  
 024-ATGTtTtTtTACACTGCACATATGTGCAACTACTGAATCACATGTGATGTGCTATATAAAATAAAT-024

6 7662 gtgcacatcatcctgccaaCcACACACCTGGCgcCAGGGGcGGGTATTGC cTactcATAA  
 11 7706 CAACACACCTGGC CAGGGGcGGGTATTGCaTgAcTaaTgTAcacATAA  
 5 33 7662 tgtacTtgctgcAttgacTcAtatataCatGCAGTgcCaATgcaaaATcTtTAATTgacTaatAgtT  
 16 7666 caoTATgycgcCAACgcctTacatACegCtgtTAGgcacATatTTTggcTtgTtTAactAAcCTAAAT  
 31 7657 TgTttaaTACACtctagtctcaACATATgtTgtcAtgcAcATATtataTTTCCTAGacAcCTTAAA  
 10 18 7626 aGgTgcgcTACAC aATTgcTgcataAcCTATAT ccactcCCTA AgtaaTAAAA  
 con tg-tatg-tacaacgccatc-a-acaaactgg-agca-satt-tata-t-cttt-cta-a--actaaaa  
 015-AGGTGGCGCTACAC AATTGCTTGCATAAATATAT CCACCTCCTA AGTAATAAAA-015  
 024-CACATATCGGCCAACGCCCTACATACCGCTGTAGGCACATATTTTGGCTTGTTTTAACTAACTAAAT-024  
 15 6 7723 ACCTGTC TTGTgttAtAcTtTaTgcAcTGTAGCCAActcTAAAJAGCATTTTTGGCTTgTAGCa  
 11 7753 ACCTGTCGTTTGT ACAaTgtTgtGgATTGCAGCCAAaggTTAAAJAGCATTTTTGGCTTGTAGCt  
 20 33 7730 TaCAcATGctTTtaggcACATAtTTTTactTTaCtttCAAAcCTTAAgTGCAGTTTTGGCTT aCa  
 16 7734 TgCATATtTGGCAtAaggTTTAAactTGTCAaggCCAAcTAAagtTCAcctAGTTCaCaATgAActg  
 31 7725 CTGCTTTTAGGCACATATTTT CTgaTTATctatATcctTgATTGCAgTgcTGGCTTTtgcacAtgt  
 18 7680 CTGCTTTTAGGCACATATTTTAgTgtTtTtctacTAAgCTAAATGCAtactTGGCTT  
 25 con --c-tttta--atataat-tagtttt-tattgct--caaa--tTaaa-gcattt-t-gctttagc-  
 BE31-XXAGGCACAXXXX-BE31 hpv16+18+33  
 BE31-XXAGGCACAXXXX-BE31 hpv16+18+33  
 015-CTGCTTTTAGGCACATATTTTAgTgtTtTtctacTAAgCTAAATGCCACTTGGCTT-(015)  
 024-TGCATATTTGGCATAAGCTTTAAACCTCTAAGGCCAACTAAATGTCAACCTAGTTCATCATGACTG-024  
 30 6 7789 GcACATTTTTtTgcCTTAcTgTtTgTgTatACAAATaCaataAAAAATGAGTAACCTAAGGTCACACACC  
 11 7818 GACATATTTTGTACCTTgTgTtTatTgCAATAcCcaAAAAATGAGTAACCTAAGGTCACACACC  
 35 33 7795 cAAAttgcTTTGTATgCcaAactATgcCTTGTAAAAgTgagtcActactcgtttTatAccaGGGTGTga  
 16 7802 TgtAAagGTTAgtctTAcATtTgTTCATTTCTAAAA CTGcAcagtGGGTGTgT  
 31 7792 TtaAAcTgcCAaggTTgtgTcaTgcATTataAATAAgTgtTatgttactcATATAATtAATtgcATat  
 18 7738 gtacaactactTtcaTgtcccaAcataTctTctaccctTtaacatgaactATAAT ATgacTaag  
 40 con --aa-attttt-tact-ttatt-tt-a-ttttaaaaaac-gtaaa-tg--tat-t-taagga-g--ta--  
 015- GTACAACTACTTTCTATGCCAACATCTGTCTACCTTAACACTGAACATAAT ATGCAATAG-015  
 024-TGTAAAGGTTAGTCATACATTGTTCATTGTAAAA CTGCACATCGGTGTGTG-024  
 45  
 50  
 55

25  
Claims for the following Contracting States : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

1. A composition useful in LCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

52

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR5 (SEQ ID Nos. 81, 82, 83 and 84) and LCR8 (SEQ ID Nos. 93, 94, 95 and 96).
3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR 6 (SEQ ID Nos. 85, 86, 87 and 88) and LCR 7 (SEQ ID Nos. 89, 90, 91 and 92).
4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:  
a composition according to any of claims 1 to 3; and  
further comprising a ligase.
5. A kit according to claim 4, wherein said ligase is thermostable.
6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising:

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences

SEQ ID No.

1 CAGATGTC TC TGTCGCGCC TAGTG.

6 GAATTAGTTA GACCATTTAA AAG.

7 GGGGAAACAC CAGAATGGAT A,

81 GCTGCAACA ACTATACATG ATATAA,

85 CTTCAC TGCA AGACATAGAA ATAA,

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences

SEQ ID No.

5 AGGTGTCAGG AAAACCAAT TTATT.

84 TGCTTGCA GT ACACACATTC TAATA,

88 TACTGTCTTG CAATATACAC AGG,

92 AATGCAAA TT CAAATACCTC GTTAA and

96 AAATCACACA ACGTTTGT TT GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences (identified by Sequence ID No.):  
1 and 5, 6 and 5, 7 and 5, 81 and 84,  
85 and 88, 89 and 92, and 93 and 96.
8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:  
a composition according to claim 6 or 7; and  
further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable.

10. A consensus oligonucleotide for hybridizing human papilloma virus types 6, 11, 16, 18, 31, 33 and 61, which oligonucleotide comprises from about 10 to about 60 nucleotides in length and is selected from the group of sequences consisting of:

SEQ ID No.

1	CAGATGTCCTC	TGTGGCGGCC	TAGTG,
5	AGGTGTCAGG	AAAACCAAAT	TTATT,
6	GAATTAGTTA	GACCATTTAA	AAG and
7	GGGAAACAC	CAGAAATGGAT	A,

and their complements.

11. A type-specific oligonucleotide for determining the presence of human papilloma virus type 16, having a sequence selected from the group consisting of:

SEQ ID No.

81	GCTGCAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTCTTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTGTGTG	ATTT and
96	AAATCACACA	ACGTTTGTG	GTAT,

and their complements.

12. A type-specific oligonucleotide for determining the presence of human papilloma virus type 18, having a sequence selected from the group consisting of: SEQ ID No.

SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTTGCAA	TTTA,
91	TTACAGAGGT	ATTGAATT	GCATT and
92	AATGCAAATT	CAATACCTC	TGTAA,

and their complements.

13. A method for determining the presence of any human papilloma virus in a test sample, comprising

a hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of claim 10, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b determining the presence of human papilloma virus by detecting the signal generated

14. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising

a hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 11, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b determining the presence of human papilloma virus by detecting the signal generated

15. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising
- hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 12, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal, and
  - determining the presence of human papilloma virus by detecting the signal generated
16. A method according to any of claims 13-15, further comprising a step of amplification prior to or concurrent with said hybridizing step
17. A method according to claim 16, wherein said amplification step comprises PCR or LCR.

#### Claims for the following Contracting States : ES

1. A composition useful in LCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

#### LCR5: SEQ ID No.

81	GCTGCAAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA;

#### LCR6: SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GCTTTGCACT	GAA,
87	CCTGTGTATA	TGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG;

#### LCR7: SEQ ID No.

89	TATATTGCAG	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGGAATTT	GCAATT,
92	AATGCAAAAT	CAAATACCTC	TGTAA; and

#### LCR8: SEQ ID No.

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATT,
96	AAATCACACA	ACGGTTTGTT	GTAT.

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR5 (SEQ ID Nos. 81, 82, 83 and 84) and LCR6 (SEQ ID Nos. 93, 94, 95 and 96).
3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR5 (SEQ ID Nos. 85, 86, 87 and 88) and LCR 7 (SEQ ID Nos. 89, 90, 91 and 92).
4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising

EP 0 477 972 B1

a composition according to any of claims 1 to 3; and further comprising a ligase

5. A kit according to claim 4, wherein said ligase is thermostable.

6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising:

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

1 CAGATGTCTC TGTGGCGGCC TAGTG.

6 GAATTAGTTA GACCATTTAA AAG.

7 GGGGAAACAC CAGAATGGAT A.

81 GCTGCAAAACA ACTATACATG ATATAA.

85 CTTCATGCA AGACATAGAA ATAA.

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

5 AGGTGTCAGG AAAACCAAAT TTATT.

84 TGCTTGCACT ACACACATTC TAATA.

88 TACTGTCTTG CAATATACAC AGG.

92 AATGCAAAAT CAAATACCTC GTAA and

96 AAATCACACA ACGTTTGTG GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences (identified by Sequence ID No.):

1 and 5, 6 and 5, 7 and 5, 81 and 84,

85 and 88, 89 and 92 and 93 and 96.

8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising a composition according to claim 6 or 7, and further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable.

10. A method for determining the presence of any human papilloma virus in a test sample, comprising

a. hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of sequences consisting of:



SEQ ID No.

1	CAGATGTCTC	TGTGGCGGCC	TAGTG.
5	AGGTGTCAGG	AAAACCAAAT	TTATT,
6	GAATTAGTTA	GACCATTTAA	AAG and
7	GGGAAACAC	CAGAATGGAT	A;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal, and

b. determining the presence of human papilloma virus by detecting the signal generated.

11. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising:

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

81	GCTGCAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT and
96	AAATCACACA	ACGGTTTGT	GTAT;

and their complements, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

12. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising:

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTICTAT	GTCTTGCACT	GAA,
87	CCTGTGTA?A	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGAATTT	GCAAT and
92	AATGCAAAAT	CAAATACCTC	TGTAA;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal, and

b. determining the presence of human papilloma virus by detecting the signal generated.

13. A method according to any of claims 10-12, further comprising a step of amplification prior to or concurrent with said hybridizing step.

14. A method according to claim 13, wherein said amplification step comprises PCR or LCR.

## Patentansprüche

Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht:

## LCR5: SEQ ID N°

81	GCTGCAAAACA	ACTATACATG	ATATAA.
82	TTATATCATG	TATAGTTGTT	TGCAGC.
83	TATTAGAATG	TGTGTACTGC	AAGCA.
84	TGCTTGCACT	ACACACATTC	TAATA.

## LCR6: SEQ ID N°

85	CTTCACTGCA	AGACATAGAA	ATAA.
86	TTATTTCTAT	GTCTTGCACT	GAA.
87	CCTGTGTATA	TTGCAAGACA	GTAT.
88	TACTGTCTTG	CAATATACAC	AGG.

## LCR7: SEQ ID N°

89	TATATTGCAA	GACAGTATTG	GAAC.
90	GTTCGAATAC	TGTCTTGCAA	TTTA.
91	TTACAGAGGT	ATTTGAATTT	GCATT.
92	AATGCAAAAT	CAATACCTC	TGTAA. und

## LCR8: SEQ ID N°

93	GTATGGAACA	ACATTAGAAC	AGCA.
94	TGCTGTTCTA	ATGTTGTTC	ATAC.
95	ATACACACAA	CCGTTGTGTG	ATTT.
96	AAATCACACA	ACGGTTTGTT	GTAT.

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR5 (SEQ ID N° 81, 82, 83 und 84) und LCR6 (SEQ ID N° 85, 86, 87 und 88).
3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 18, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR6 (SEQ ID N° 85, 86, 87 und 88) und LCR7 (SEQ ID N° 89, 90, 91 und 92).
4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.
5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.
6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" Polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

# EP 0 477 972 B1

einen ersten Nukleinsäureprimer, der zur Richtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht

5  
1 CAGATGTC TC TGTGGCGGCC TAGTG.  
6 GAATTAGTTA GACCATTTAA AAG,  
10 7 GGGGAAACAC CAGAATGGAT A,  
81 GCTGCAAAACA ACTATACATG ATATAA,  
85 CTTCACTGCA AGACATAGAA ATAA,  
89 TATATTGCAA GACAGTATTG GAAC und  
15 93 GTATGGAACA ACATTAGAAC AGCA; und

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht.

## 20 SEQ ID Nr

25 5 AGGTGTCAGG AAAACCAAAT TTATT,  
84 TGTCTTCAGT ACACACATTC TAATA,  
88 TACTGTCCTG CAATATACAC AGG,  
92 AATGCAAATT CAAATACCTC GTAA und  
30 96 AAATCACACA ACGTTTGTI GTAT;

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer.

- 35 7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:  
1 und 5, 6 und 5, 7 und 5, 81 und 84,  
85 und 88, 89 und 92, und 93 und 96
- 40 8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach Anspruch 6 oder 7 und des weiteren eine Polymerase
- 45 9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist
10. Consensus-Oligonukleotid zur Hybridisierung der humanen papillomaviren Typ 6, 11, 16, 18, 31, 33 und 61, wobei das Oligonukleotid ungefähr 10 bis ungefähr 60 Oligonukleotide lang ist und aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht

## 50 SEQ ID Nr

51 CAGATGTC TC TGTGGCGGCC TAGTG.  
52 AGGTGTCAGG AAAACCAAAT TTATT,  
55 6 GAATTAGTTA GACCATTTAA AAG und  
7 GGGGAAACAC CAGAATGGAT A;

und aus deren Komplementen.

11. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

SEQ ID Nr

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTTG	ATT und
96	AAATCACACA	ACGGTTTGT	GTAT;

und aus deren Komplementen.

12. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

SEQ ID Nr

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT und
92	AATGCAAATT	CAAAATACCT	.TGTA;

und aus deren Komplementen.

13. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe nach Anspruch 10 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und  
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

14. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 11 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und  
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

15. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 12 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

16. Verfahren nach einem der Ansprüche 13-15, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder in Konkurrenz mit dem Hybridisierungsschritt stattfindet

17. Verfahren nach Anspruch 16, worin der Vervielfachungsschritt PCR oder LCR umfaßt

#### Patentansprüche für folgenden Vertragsstaat : ES

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht:

#### LCR5: SEQ ID N°

```
81 GCTGCAAAACA ACTATACATG ATATAA,
82 TTATATCATG TATAGTTGTT TGCAGC,
83 TATTAGAATG TGTGTACTGC AAGCA,
84 TGCTTGCACT ACACACATTC TAATA;
```

#### LCR6: SEQ ID N°

```
85 CTTCACCTGCA AGACATAGAA ATAA,
86 TTATTTCTAT GTCTTGCACT GAA,
87 CCTGTGTATA TTGCAAGACA GTAT,
88 TACTGTCTTG CAATATACAC AGG;
```

#### LCR7: SEQ ID N°

```
89 TATATTGCAA GACAGTATTG GAAC,
90 GTTCCAATAC TGTCTTGCAA TTTA,
91 TTACAGAGGT ATTTGAATTT GCATT,
92 AATGCAAAAT CAAATACCTC TGTA.
```

#### LCR8: SEQ ID N°

```
93 GTATGGAACA ACATTAGAAC AGCA, und
94 TGCTGTTCTA ATGTTGTTCC ATAC,
95 ATACAACAAA CCGTTGTGTG ATTT,
96 AAATCACACA ACGGTTTGTT GTAT.
```

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR5 (SEQ ID N° 81, 82, 83 und 84) und LCR8 (SEQ ID N° 93, 94, 95 und 96)

3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus TYP 18, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die Sondensätze aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR6 (SEQ ID N° 85, 86, 87 und 88) und LCR7 (SEQ ID N° 89, 90, 91 und 92).

4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.

5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.

6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

einen ersten Nukleinsäureprimer, der zur Richtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

1	CAGATGTCCTC	TGTGGCGGCC	TAGTG.
6	GAATTAGTTA	GACCATTTAA	AAG,
7	GGGAAACAC	CAGAAATGGAT	A,
81	GCTGCAACA	ACTATACATG	ATATAA,
85	CTTCACTGCA	AGACATAGAA	ATAA,
89	TATATTGCAA	GACAGTATTG	GAAC und
93	GTATGGAACA	ACATTAGAAC	AGCA; und

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

5	AGGTGTCAGG	AAAACCAAT	TTATT,
84	TGCTTGCACT	ACACACATTC	TAATA,
88	TACTGTCTTG	CAATATACAC	AGG,
92	AATGCAAAAT	CAATACCTC	TGTAA und
96	AAATCACACA	ACGGTTTGTT	GTAT;

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer

7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:

1 und 5, 6 und 5, 7 und 5, 81 und 84,  
85 und 88, 89 und 92, und 93 und 96

8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

eine Zusammensetzung nach Anspruch 6 oder 7, und des weiteren eine Polymerase.

9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist.

10. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht:

SEO ID Nr

5

10

```

1   CAGATGTC TC TGTGGCGGCC TAGTG,
5   AGGTGTCAGG AAAACCAAAT TTATT
6   GAATTAGT TA GACCATTTAA AAG und
7   GGGGAAACAC CAGAAATGGAT A,

```

15

und aus deren Komplementen,

wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

20

11. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes umfaßt:

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht.

25

SEO ID Nr

30

35

40

```

81  GCTGCAACA ACTATACATG ATATAA,
82  TTATATCATG TATAGTTGTT TGCAGC,
83  TATTAGAATG TGTGTACTGC AAGCA,
84  TGCTTGCACT ACACACATTC TAATA,
93  GTATGGAACA ACATTAGAAC AGCA,
94  TGCTGTTCTA ATGTTGTTCC ATAC,
95  ATACAACAAA CCGTTGTGTG ATTT und
96  AAATCACACA ACGTTTGTGT GTAT;

```

und aus deren Komplementen,

wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird

45

12. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes umfaßt:

50

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht:

55

SEQ ID Nr

5                   85   CTTCACTGCA AGACATAGAA ATAA.  
                   86   TTATTTCTAT GTCTTGCAGT GAA,  
                   87   CCTGTGTATA TTGCAAGACA GTAT,  
                   88   TACTGTCTTG CAATATACAC AGG,  
                   89   TATATTGCAA GACAGTATTG GAAC,  
                   90   GTTCGAATAC TGTCTTGCAA TTTA,  
 10                91   TTACAGAGGT ATTGAATTT GCATT und  
                   92   AATGCAAAAT CAAATACCTC TGTAAT;

15                   und aus deren Komplementen,  
                   wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nach-  
                   weisbaren Signals befähigt ist, und  
                   b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

20   13. Verfahren nach einem der Ansprüche 10-12, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder  
                   in Konkurrenz mit dem Hybridisierungsschritt stattfindet

14. Verfahren nach Anspruch 13, worin der vervielfachungsschritt PCR oder LCR umfaßt

25   **Revendications**

**Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE**

30   1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser,  
                   ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes  
                   étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants.

35   LCR5 : n° d'identification  
                   81   GCTGCAACA ACTATACATG ATATAA,  
                   82   TTATATCATG TATAGTTGTT TGCAGC,  
                   83   TATTAGAATG TGTGTACTGC AAGCA,  
 40                84   TGCTTGCAGT ACACACATTC TAATA;

45   LCR6 : n° d'identification  
                   85   CTTCACTGCA AGACATAGAA ATAA,  
                   86   TTATTTCTAT GTCTTGCAGT GAA,  
                   87   CCTGTGTATA TTGCAAGACA GTAT,  
 50                88   TACTGTCTTG CAATATACAC AGG;

55



## LCR7 : n° d'identification

89	TATATTGCAA	GACAGTATTG	GAAC
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGGAATTT	GCATT,
92	AATGCAAATT	CAAATACCTC	TGTAA ; et

## LCR8 : n° d'identification

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT,
96	AAATCACACA	ACGGTTTGTT	GTAT.

2. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 16 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR5 (n° d'identification 81, 82, 83 et 84) et LCR8 (n° d'identification 93, 94, 95 et 96).

3. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 18 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR6 (n° d'identification 85, 86, 87 et 88) et LCR7 (n° d'identification 89, 90, 91 et 92).

4. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant : une composition selon l'une quelconque des revendications 1 à 3, et en outre une ligase.

5. Kit selon la revendication 4, dans lequel ladite ligase est thermostable.

6. Composition utile dans la PCR pour amplifier l'ADN de virus du papillome humain présent dans un échantillon à doser, ladite composition comprenant :

une première amorce d'acide nucléique de direction sens, capable de s'hybrider au brin antisens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

## N° d'identification

1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
6	GAATTAGTTA	GACCATTTAA	AAG,
7	GGGGAACAC	CAGAATGGAT	A,
81	GCTGCAACAA	ACTATACATG	ATATAA,
85	CTTCACTGCA	AGACATAGAA	ATAA,
89	TATATTGCAA	GACAGTATTG	GAAC et
93	GTATGGAACA	ACATTAGAAC	AGCA ; et

une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN

de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

5	N° d'identification	5	AGGTGTCAGG AAAACCAAAT TTATT,
		84	TGCTTGCACT ACACACATTC TAATA,
		88	TACTGTCITG CAATATACAC AGG,
10		92	AATGCAAATT CAAATACCTC TGTA et
		96	AAATCACACA ACGGTTTGT GTAT ;

pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces.

- 15 7. Composition selon la revendication 6, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :  
1 et 5, 6 et 5, 7 et 5, 81 et 84,  
20 85 et 88, 89 et 92, et 93 et 96.
8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon la revendication 6 ou 7 et en outre une polymérase.
- 25 9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable.
10. Oligonucléotide consensus pour hybridation du virus du papillome humain des types 6, 11, 16, 18, 31, 33 et 61, lequel oligonucléotide a d'environ 10 à environ 60 nucléotides de long et est sélectionné dans le groupe de séquences constitué par :

20	N° d'identification	1	CAGATGCTC	TGTGGCGGCC	TAGTG,
		5	AGGTGTCAGG	AAAACCAAAT	TTATT,
35		6	GAATTAGTTA	GACCATTTAA	AAG et
		7	GGGAAACAC	CAGAATGGAT	A ;

et leurs compléments

- 40 11. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 16, ayant une séquence sélectionnée dans le groupe constitué par :

45	N° d'identification	81	GCTGCAACA	ACTATACATG	ATATAA,
		82	TTATATCATG	TATAGTTGTT	TGCAGC,
		83	TATTAGAATG	TGTGTACTGC	AAGCA,
		84	TGCTTGCACT	ACACACATTC	TAATA,
		93	GTATGGAACA	ACATTAGAAC	AGCA,
50		94	TGCTGTTCTA	ATGTTGTTC	ATAC,
		95	ATACAACAAA	CCGTTGTGTG	ATTT et
		96	AAATCACACA	ACGGTTTGT	GTAT ;

et leurs compléments

- 55 12. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 18, ayant une séquence sélectionnée dans le groupe constitué par :

## N° d'identification

	85	CTTCACTGCA	AGACATAGAA	ATAA,
5	86	TTATTCTAT	GTCTGCACT	GAA,
	87	CCTGTGTATA	TTGCAAGACA	GTAT,
	88	TACTGTCTTG	CAATATACAC	AGG,
	89	TATATTGCAA	GACAGTATTG	GAAC,
10	90	GTTCCAATAC	TGTCITGCAA	TTTA,
	91	TTACAGAGGT	ATTTGAATT	GCATT et
	92	AATGCAAAIT	CAAATACCTC	TGTAA ;

et leurs compléments.

13. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe selon la revendication 10, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

14. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 11, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

15. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 12, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

16. Procédé selon une quelconque des revendications 13 à 15, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.

17. Procédé selon la revendication 16, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

Revendications pour l'Etat contractant suivant : ES

1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants.

## LCRS : n° d'identification

	81	GCTGCAAACA	ACTATACATG	ATATAA,
	82	TTATATCATG	TATAGTTGTT	TGCAGC,
55	83	TATTAGAATG	TGTGTACTGC	AAGCA,
	84	TGCTTGCAGT	ACACACATTC	TAATA ;

## LCR6 : n° d'identification

85	CTTCACTGCA	AGACATAGAA ATAA,
86	TTATTTCTAT	GTCTTGCACT GAA,
87	CCTGTGTATA	TTGCAAGACA GTAT,
88	TACTGTCTTG	CAATATACAC AGG ;

## LCR7 : n° d'identification

89	TATATTGCAA	GACAGTATTG GAAC
90	GTTCCAATAC	TGTCITGCAA TTTA,
91	TTACAGAGGT	ATTTGAATTT GCATT,
92	AATGCAAATT	CAAATACCTC TGTA ; et

## LCR8 : n° d'identification

93	GTATGGAACA	ACATTAGAAC AGCA,
94	TGCTGTCTA	ATGTTGTTC ATAC,
95	ATACAACAAA	CCGTTGTGTG ATTT,
96	AAATCACACA	ACGTTTGTG GTAT.

2. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 16 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR5 (n° d'identification 81, 82, 83 et 84) et LCR8 (n° d'identification 93, 94, 95 et 96).

3. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 18 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

LCR6 (n° d'identification 85, 86, 87 et 88) et LCR7 (n° d'identification 89, 90, 91 et 92).

4. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon l'une quelconque des revendications 1 à 3, et en outre une ligase

5. Kit selon la revendication 4, dans lequel ladite ligase est thermostable.

6. Composition utile dans la PCR pour amplifier l'ADN de virus du papillome humain présent dans un échantillon à doser, ladite composition comprenant :

une première amorce d'acide nucléique de direction sens, capable de s'hybrider au brin antisens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

## N° d'identification

5	1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
	6	GAATTAGTTA	GACCATTTAA	AAG,
	7	GGGGAAACAC	CAGAATGGAT	A,
10	81	GCTGCAAAAC	ACTATACATG	ATATAA,
	85	CTTCACTGCA	AGACATAGAA	ATAA,
	89	TATATTGCAA	GACAGTATTG	GAAC et
	93	GTATGGAACA	ACATTAGAAC	AGCA ; et

15 une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

## N° d'identification

20	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	84	TGCTTGCACT	ACACACATTC	TAATA,
	88	TACTGTCTTG	CAATATACAC	AGG,
25	92	AATGCAAATT	CAAATACCTC	TGTAA et
	96	AAATCACACA	ACGGTTTGTT	GTAT ;

30 pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces.

7. Composition selon la revendication 6, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :  
1 et 5, 6 et 5, 7 et 5, 81 et 84,  
35 85 et 88, 89 et 92, et 93 et 96

8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant une composition selon la revendication 6 ou 7 et en outre une polymérase

40 9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable

10. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser, comprenant :

45 a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe de séquences constitué par :

## N° d'identification

50	1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	6	GAATTAGTTA	GACCATTTAA	AAG et
	7	GGGGAAACAC	CAGAATGGAT	A ;

55 et leurs compléments, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

11. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

**N° d'identification**

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTGTG	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTCTA	ATGTTGTTC	ATAC,
95	ATACAACAAA	COGTTGTGTG	ATTT et
96	AAATCACACA	ACGGTTGTG	GTAT ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

12. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

**N° d'identification**

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTGGAATTT	GCATT et
92	AATGCAAATT	CAAATACCTC	TGTAA ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

13. Procédé selon une quelconque des revendications 10 à 12, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.

14. Procédé selon la revendication 13, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

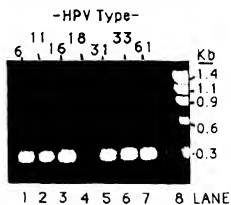


FIG. 1

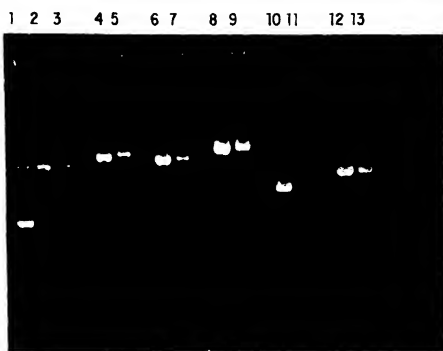


FIG. 2



FIG. 3



FIG. 4

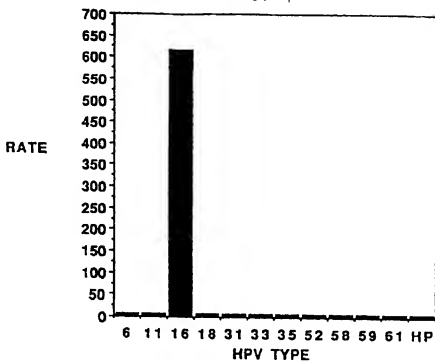


FIG. 5

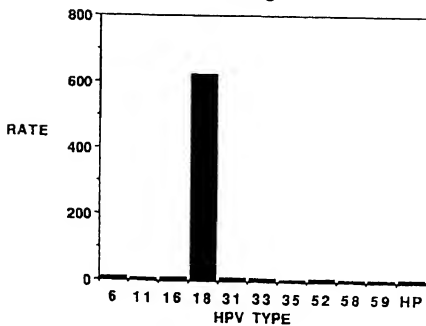


FIG. 6

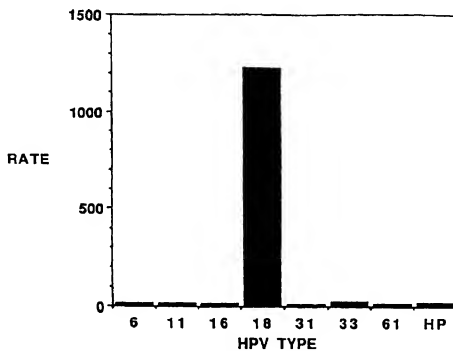


FIG. 7

